8-2012

Generality of Treatment Effects: Evaluating Elementary-Aged Students' Abilities to Generalize and Maintain Fluency Gains of a Performance Feedback Writing Intervention

Bridget Hier
Syracuse University

Abstract

Although writing ability is a skill that has been argued to be equally important as reading skills in the development of early literacy and is necessary for academic success (Berninger et al., 2006; Graham, MacArthur, & Fitzgerald, 2007), national estimates of students' writing ability in the United States indicate that in 2002, 72% of elementary-aged students were unable to write with grade-level proficiency (Persky, Daane, & Jin, 2003). This finding presents a clear need for empirical, evidence-based interventions that aim to improve students' writing skills, and performance feedback is one type of intervention that has been shown to do so (Eckert et al., 2006). However, no study to date has examined the generalization and maintenance of writing fluency gains that have been developed as a result of performance feedback interventions. The primary goal of this study was to determine whether 51 third-grade students assigned to a performance feedback intervention condition demonstrated evidence of greater (a) writing fluency gains, (b) generalization of writing fluency, and (c) maintenance of writing fluency in comparison to 52 students assigned to a practice-only control condition. Results revealed that although students assigned to the performance feedback condition demonstrated significantly greater writing fluency growth during the course of the intervention than students assigned to the practice-only condition, evidence for maintenance and generalization of intervention effects was limited. These findings suggest that, in isolation, performance feedback may produce short-term desired effects on students' writing fluency growth, but that explicit programming of generality may be required to produce long-term achievement gains.

Keywords: academic intervention, writing, performance feedback, generality, generalization, maintenance

GENERALITY OF TREATMENT EFFECTS: EVALUATING ELEMENTARY-AGED STUDENTS' ABILITIES TO GENERALIZE AND MAINTAIN FLUENCY GAINS OF A PERFORMANCE FEEDBACK WRITING INTERVENTION

by

Bridget Hier

B.A., North Carolina State University, 2009

THESIS

Submitted in partial fulfillment of the requirements for the degree of
Master of Science in *Psychology*.

Syracuse University
August 2012

TABLE OF CONTENTS

Generality of Treatment Effects: Evaluating Elementary-Aged Students' Ability to Generalize and Maintain Fluency Gains of a Performance Feedback Writing Intervention

Writing ability, a skill that has been argued to be equally important as reading skills in the development of early literacy, is necessary for academic success (Berninger et al., 2006; Graham, MacArthur, & Fitzgerald, 2007). It is a communication tool that allows for learning, persuasion, self-expression, and knowledge refinement (Graham et al., 2007).Writing continues to be a necessary skill after the completion of elementary and secondary school. Students' writing samples are often used in university entrance decisions, and many university students' writing skills continue to be evaluated throughout their academic careers. Despite this emphasis on writing, university faculty estimated that as many as 50% of high school graduates are unprepared to write at the college level (Achieve, Inc., 2005). Writing is also an important skill in the workplace, where many employees are expected to write reports and prepare presentations. Yet, employers estimated that 38% of high school graduates' writing quality is not sufficient for the workplace (Achieve, Inc., 2005).

The development of effective writing skills is an essential aspect of elementary and secondary education (National Commission on Writing, 2003). Because writing can transmit knowledge, it is used as a primary instrument for evaluating academic competence (Graham et al., 2007). Failure to gain the writing abilities necessary for satisfactory schoolwork can put children at risk for behavioral problems, chronic school failure, and school dropout (Berninger et al., 2006). National estimates of students' writing ability in the United States suggest a substantial need for improvement. The National Assessment of Educational Progress, published by the U.S. Department of Education, reports on national writing data from fourth-, eighth-, and

twelfth-grade students. Results from these studies revealed that in 2002, 72% of fourth-grade students were unable to write at the Proficient level (i.e., a level that displays mastery of grade-level expectations; Persky, Daane, & Jin, 2003), and in 2007, 67% of eighth-, and 76% of twelfth-grade students were unable to write at the Proficient level (Salahu-Din, Persky, & Miller, 2008). These writing difficulties were exacerbated for students of minority backgrounds. Among the fourth-grade students assessed, 86% of Black students, 83% of Hispanic students, and 86% of American Indian/Alaska Native students performed below the Proficient level. Although White students and Asian/Pacific Islander students demonstrated higher levels of proficiency, a substantial percentage (i.e., 67% and 59%, respectively) were unable to reach the Proficient level (Persky et al., 2003). Furthermore, 88% of fourth-grade students of low socioeconomic status (i.e., eligible for free/reduced lunch price) were unable to write at the Proficient level.

Given the substantial percentages of students throughout the nation writing at a level of inadequacy, research efforts should focus on methods designed to improve writing skill. Furthermore, efforts should be made to ensure gains achieved from writing interventions are able to be generalized and maintained. The purpose of this study was to examine the effects of a performance feedback intervention on elementary-aged students' writing fluency (i.e., the ability to write with speed and accuracy). An additional aim of the study was to determine whether students are able to generalize any gains in writing fluency to different writing tasks. A final aim of the study was to examine whether students were able to maintain any gains in writing fluency over a period of 2 to 6 weeks.

Theoretical Conceptualizations of Writing

Hayes and Flower (1980) proposed an influential model of writing that emphasized cognitive processes involved in the writing process. This model was comprised of three

components: (a) planning, (b) translating, and (c) reviewing/revising the written product. Prior to text generation, Hayes and Flower proposed that children engage in a period of planning. During the planning component, children are thought to engage in idea generation, organizing, and goal-setting regarding their written work. Following the planning component, it was proposed that children engage in translation, which was conceptualized as converting language representations into orthographic text. Translation was hypothesized to rely heavily on the transfer of ideas in working memory. Finally, it was proposed that children review, evaluate, and revise their written products (see Figure 1). Hayes and Flower theorized that unlike the translating component, both the planning and reviewing components were comprised of multiple sub-component processes.

Berninger et al. (1997) argued that translation plays a significant role in the writing process for children developing these emergent skills, and that the Hayes and Flower (1980) model thus was more applicable to the adult writing process. Because planning and revision require higher-level processes, children who have not yet mastered the lower-level process of text production may have difficulty engaging in these components. Berninger and colleagues thus proposed a model that emphasized the importance of translation in emerging writers (Abbott & Berninger, 1993). Further, Berninger, Yates et al. (1992) proposed dividing the component of translation into two sub-components: (a) text generation and (b) transcription (see Figure 2). Text generation involves transforming ideas into linguistic representations in working memory. Once linguistic representations are formed, children are thought to engage in transcription, which transfers these linguistic representations into motor output. It is this motor-based sub-component of translation (i.e., transcription) that is directly related to the mechanics of writing (e.g., spelling, punctuation, and grammar), handwriting, and compositional fluency (Berninger et al., 2006). Two transcription skills in particular (i.e., spelling and handwriting) have been shown to

be related to the length and quality of compositions (Graham, Berninger, Abbott, Abbott, & Whitaker, 1997). Research supports the idea that these two sub-components play a significant role in the writing process for primary-grade children (Berninger, Cartwright, Yates, Swanson, & Abbott, 1994) and intermediate-grade children (Berninger et al., 2002).

Developmentally-Appropriate Writing Practices

Relatively few studies examining the developmental appropriateness of writing practices exist. Abbott and Berninger (1993) suggested transcription should be a focus of primary-level teaching practices, as it is considered to be a fundamental skill that is relied upon by emerging writers. Research efforts have focused on handwriting instruction, a component of transcription, for children in kindergarten and first grade (Berninger et al., 1997, 2006). Theoretically, children for whom handwriting is an automatic skill will have the ability to focus cognitive resources on higher-level writing processes (e.g., text generation, planning, reviewing; Berninger et al., 1997, 2006). Evidence suggests that handwriting does in fact account for a significant proportion of variance in children's writing fluency (r = .68 to .99; Abbott & Berninger, 1993; Graham et al., 1997). Furthermore, students' writing outcomes were shown to improve as a result of intensive instruction in handwriting (Berninger et al., 1997, 2006).

For students who have mastered the transcription component of writing, it may be more developmentally appropriate to focus instruction on the planning and reviewing components, which are thought to be higher-level writing processes. Instructional practices targeted toward the development of planning and reviewing have focused on strategy instruction for both elementary and secondary students. Troia (2002) suggested that effective strategy instruction should include brainstorming words and ideas, generating and organizing content with prompts, and setting and planning goals. A well-researched strategy instructional model, Self-Regulated

Strategy Development (Harris & Graham, 1999), includes five stages: (a) develop and activate background knowledge (i.e., learn and understand basic parts of a story), (b) discussion between the teacher and student regarding instructional goals, (c) teacher-facilitated modeling of goal-setting, story composition, and reviewing through "thinking aloud," (d) student memorization of strategy steps, and (e) support from the teacher to help the student use the steps to plan a story (Harris, Santangelo, & Graham, 2008).

A meta-analysis examining Self-Regulated Strategy Development and other forms of strategy instruction was conducted by Graham (2006). He reviewed 39 studies in which participants ranged from first to twelfth grade students with and without learning disabilities. Overall, students who received some form of strategy instruction improved in their writing quality ($\Delta = 1.16$), genre elements of writing (e.g., basic components of a composition, main characters, location, time frame; $\Delta = 1.88$), and revising ($\Delta = .90$). Despite these large gains, the mechanics of students' writing did not seem to improve to a large degree ($\Delta = .30$). Although students with learning disabilities experienced improvement in their composition length ($\Delta = 1.39$), good writers' composition length decreased as a result of strategy instruction ($\Delta = -.02$). Overall, strategy instruction appears to be beneficial for a wide range of students; however, beginning writers who are not fluent in the transcription process will not likely benefit from strategy instruction (Abbott & Berninger, 1993). Furthermore, because composition length and mechanics do not improve as a result of strategy instruction, it is not an ideal instructional practice for improving children's writing fluency.

Current State of Writing Instruction

Given the data suggesting that a majority of children in the United States are unable to write with grade-level proficiency (Salahu-Din et al., 2008), it has been suggested that students

experience inadequate writing instruction in the classroom (Troia, 2002). Despite an improvement in writing instruction from 10 to 15 years ago (Troia, 2002), research suggests current instructional practices employed in elementary and secondary school settings are still inadequate (Cutler & Graham, 2008; National Commission on Writing, 2003). It has also been suggested that further improvements in writing instruction should be seriously considered, particularly at the elementary school level (Cutler & Graham, 2008; Graham, Harris, Fink-Chorzempa, & MacArthur, 2003). One recent study by Graham et al. (2003) found that of a national teacher sample, only 11% reported adapting their teaching methods (e.g., providing additional conferencing, re-teaching material) for their elementary-aged students struggling with basic writing skills. Furthermore, this study revealed that teachers reportedly devoted an average of only four hours per week to writing instruction, and that their students received an average of only three hours per week (i.e., 36 minutes per day) of writing practice in the classroom. These findings demonstrate the need for more classroom writing practice, as was recommended by the National Commission on Writing (2003).

Recently, Cutler and Graham (2008) conducted a study that examined the writing instructional practices utilized by a national, random sample of elementary grade teachers. A large amount of variability was found between teachers in terms of instructional approaches, with the majority of teachers (65%) reported using their own instructional strategies to teach writing. Of those teachers who reported using commercially-available writing programs, more than 100 different programs were identified. Another important finding reported in this study related to instructional practices in handwriting and spelling. Despite Berninger et al.'s (2006) recommendation to increase instructional time spent on handwriting and spelling for beginning writers, Cutler and Graham (2008) found that teachers reported providing an average of only 46

minutes per week of direct instruction in handwriting (9.2 minutes per day), and 74 minutes per week of direct instruction in spelling (14.8 minutes per day). Furthermore, the sample of teachers in Cutler and Graham's study reported that their students were only given approximately 21 minutes per day to practice writing connected text. This estimate is even lower than that found by Graham et al. (2003), again suggesting the need for greater daily writing practice in the classroom.

Writing Fluency

The development of writing fluency is a critical, fundamental skill that contributes to children's writing ability. Developing automaticity and proficiency in transcription (i.e., writing fluency) is particularly important for elementary-aged students, as the ability to write fluently allows children to expend fewer cognitive resources on basic writing components. Thus, the ability to write fluently may allow more cognitive resources to be utilized for other, higher-level writing components such as composition planning and content knowledge (Graham et al., 1997). Given that fluency is assumed to be a fundamental skill in the writing process, it has been argued that the development of this skill should be targeted in the elementary grades (Abbott & Berninger, 1993; Berninger et al., 2006).

Writing fluency is correlated with a number of other writing indices, including writing quality (Deno, Mirkin, & Marston, 1980; Van Houten, Morrison, Jarvis, & McDonald, 1974), criterion and standardized measures of writing achievement (McMaster & Espin, 2007; Powell-Smith & Shinn, 2004), and even post-secondary educational success (Calfee & Miller, 2007). Writing fluency is typically assessed using Curriculum-Based Measurement in Written Expression (CBM-WE), which provides students with a short story-starter prompt and allows students 3 minutes to write a narrative. From the narrative, a number of measures can be derived,

the most common of which are total number of words written, total number of correctly spelled words, and total number of correct writing sequences (Espin et al., 2000). Other measures that can be scored from CBM-WE narratives include total number of letters written, number of complete sentences, number of words in complete sentences, and number of correct punctuation marks. The reliability and validity of CBM-WE as an indicator of writing fluency among elementary-aged students has been supported by an extensive number of studies (Powell-Smith & Shinn, 2004).

According to the National Writing Project (2003), writing fluency is an essential skill that should be established in the elementary grades to develop writing competence. Despite this, evidence by Graham et al. (2003) and Cutler and Graham (2008) suggest that elementary-aged students do not spend a sufficient amount of time practicing writing in the classroom. A lack of opportunity to establish writing fluency skills in the elementary grades is particularly problematic because most school curricula do not continue formal writing instruction after the elementary grade levels (Smith, 2004). These instructional practices may not be optimal for the majority of students, as 72% of fourth-grade students were unable to write at the Proficient level in 2002 (Persky et al., 2003), and 43% of referrals for special education services were for students experiencing writing difficulties (Bramlett, Murphy, Johnson, & Wallingsford, 2002). Although a number of interventions targeting handwriting and strategy use have been empirically examined (Berninger et al., 1997, 2006; Graham, 2006), few interventions targeting writing fluency have been tested for effectiveness. One intervention that has been shown to reliably improve writing fluency skills in elementary-aged students is the use of performance feedback.

Performance Feedback Interventions

Performance feedback can be conceptualized as "information provided by an agent (e.g., teacher, peer, book, parent, self, experience) regarding aspects of one's performance or understanding" (Hattie & Timperley, 2007, p. 81). According to these researchers, upon receiving feedback individuals utilize cognitive and affective processes to reduce the gap between what they currently understand and what is aimed to be understood. For example, receiving feedback regarding test performance may confirm that a student's response was incorrect (i.e., cognitive process) and prompt the individual to put forth greater effort on future exams (i.e., affective process).

Interventions designed to improve academic skills through the use of performance feedback have been effectively implemented in a number of content areas including writing, reading, and mathematics (Eckert et al., 2006). Thorndike (1931) originally examined performance feedback by explicitly informing adults whether their guesses to word-meaning associations were correct or incorrect. Participants who received this type of performance feedback were more likely to make correct associations following the feedback, as opposed to a group of participants who did not receive feedback regarding their performance. As a result, performance feedback interventions are conceptualized as cognitive-behavioral, as they have the ability to manipulate one's cognitions as well as behavior. Locke, Shaw, Saari, and Latham (1981) further argued the cognitive-behavioral nature of performance feedback, stating that individuals must understand and apply the new information gained from feedback, thus changing individuals' thoughts about subsequent behavior. A cognitive-behavioral mechanism also appears to be utilized by young children acquiring academic skills through performance feedback interventions. Information gained from feedback regarding performance on academic tasks can not only modify students' thoughts about their performance, but it can also serve as a

motivational stimulus to reinforce behavior. Indeed, teachers and researchers alike report the positive impacts of performance feedback on academic skills (Moxley, Lutz, Ahlborn, Boley, & Armstrong, 1995).

Hattie and Timperley (2007) reviewed 74 meta-analyses of performance feedback studies. Considerable variability in effect sizes suggested that some aspects of performance feedback may be more powerful than others. Specifically, higher effect sizes were found for feedback that included information about the task at hand and how to perform it more effectively (range, .94 to 1.10). Feedback that included praise, extrinsic rewards, and punishment typically were associated with lower effect sizes (range, .14 to .31).

A number of single case research studies have examined the effect of performance feedback on the writing abilities of elementary students acquiring basic skills. Van Houten et al. (1974) used performance feedback as a component of an intervention package that was effective in increasing students' writing fluency. During the baseline phase of a reversal design, participants (i.e., 21 second- and 34 fifth-grade students) were instructed to write as many words as they could about a topic provided by their teacher. A timing and feedback condition was then employed, where the students were (a) told they had 10 minutes to write a composition, (b) instructed to score their composition by counting the total number of words written, (c) able to see each student's highest score on a publically posted chart in the classroom, and (d) instructed to attempt to beat their own high score. For both second- and fifth-grade students, writing fluency increased substantially with this packaged intervention. Unfortunately, due to the packaged nature of the intervention, it is impossible to evaluate the effects of performance feedback in isolation. Furthermore, these researchers did not examine students' ability to generalize and maintain their writing fluency gains.

Van Houten, Hill, and Parsons (1975) examined the contribution of performance feedback alone on elementary students' writing fluency. A reversal design was used to individually introduce self-scored performance feedback, public posting of scores, and verbal teacher praise. It was found that the provision of performance feedback alone doubled the writing rate of students in comparison to their baseline levels. Combining the elements of performance feedback and public posting of scores increased writing rates by approximately 2.2 words per minute. Introducing teacher praise increased the writing rate of one classroom by an additional 2.2 words per minute, but had little effect on the writing rate of the other classroom. It was concluded that each component (i.e., self-scored performance feedback, public posting of scores, and teacher praise) contributed to the effectiveness of the writing intervention package, and that the combination of performance feedback and public posting produced the greatest increase in students' writing rates. Although this study addressed one of Van Houten et al.'s (1974) limitations by examining performance feedback in isolation, it again failed to examine maintenance and generalization of treatment effects.

Van Houten (1979) examined students' ability to generalize their academic performance following participation in a performance feedback intervention. Sixty second- through fourth-grade students received a combination of self-scored total words written (i.e., performance feedback) and publically posted high scores. Following the performance feedback intervention, the students' writing rate increased by approximately four words per minute. However, when contingencies were not explicitly stated (i.e., students were not told they would receive feedback), students' writing rates still increased by approximately 2.2 words per minute over the baseline phase. This suggests that students were able to generalize effects of the performance feedback intervention to conditions that differed from that of the intervention (i.e., conditions in

which no contingencies were explicitly stated). However, performance feedback was not examined in isolation in this study.

Rather than examining performance feedback as a piece of a larger writing package, Harris, Graham, Reid, McElroy, and Hamby (1994) used a multiple baseline design across participants to determine the effect of performance feedback in isolation on writing performance and on-task behavior of four fifth- and sixth-grade students with learning disabilities. Students were given 15 minutes to write a story in response to a picture prompt, and then self-score their composition for the total number of words written. This performance feedback produced a substantial increase in both total words written (baseline $M = 50.25$; intervention $M = 109.50$) and writing quality as measured on an 8-point scale (baseline $M = 2.52$; intervention $M = 4.38$). Although these results support the use of performance feedback as a tool to improve the writing fluency of elementary-aged students with learning disabilities, one cannot generalize these findings to a typically-developing population. Furthermore, because maintenance and generalization of writing fluency was not examined, one cannot infer whether this type of intervention results in durability of effects over time and transfer of effects across situations.

Eckert and colleagues (2006) recently reported on their empirical research examining the effects of performance feedback in isolation on the writing fluency outcomes of typically-developing third-grade students. In these studies, students assigned to a performance feedback condition received a numeric indicator of their progress from the previous session accompanied by an upward-facing arrow, which indicated improvement in their performance (i.e., total number of words written, total number of sentences written, correct writing sequences), or a downward-facing arrow, which indicated decline. Following this performance feedback, students

were given 3 minutes to write a composition based on a story-starter prompt (e.g., “I never dreamed that door in my bedroom would lead to…”).

In their first study, Eckert and colleagues (2006) randomly assigned 50 third-grade students to either a control condition or a performance feedback condition. Over a period of eight weeks, students in the control condition received weekly writing practice without the use of performance feedback, and students in the performance feedback condition received a weekly intervention as described above. Writing progress was determined by calculating calendar day slope estimates for each participant. Results indicated that students in the performance feedback condition demonstrated a statistically significantly greater increase in writing fluency outcomes (d = .65, CI = +.25 to +1.06) and spelling (d = .74, CI = +.33 to +1.14) than those students in the control condition.

Eckert and colleagues (2006) conducted a second study examining the extent to which varying amounts of feedback affected the performance of 42 third-grade students from three different classrooms. Each classroom was randomly assigned to one of three conditions: (a) control, in which students received practice without feedback three times per week, (b) performance feedback once per week, and (c) performance feedback three times per week. Students in all conditions received three writing sessions per week over six weeks; however, those students in the once-per-week condition only received feedback during the last session of the week, while students in the three-times-per-week condition received feedback during each session. Results indicated that compared to students in the control condition, students receiving performance feedback, regardless of how often, made statistically significantly greater gains in writing fluency outcomes (i.e., total number of words written; d = .61, CI = +0.19 to +1.40).

However, the frequency of performance feedback (i.e., once per week or three times per week) did not have a significant effect on writing outcomes.

The research findings reported by Eckert et al. (2006) suggest that performance feedback can be effectively used to increase writing fluency outcomes among typically-developing elementary-aged students. However, these studies did not examine the extent to which writing fluency gains were generalized to different writing situations and maintained over time. Furthermore, these studies did not employ multilevel linear modeling, the recommended technique for analyzing classroom intervention data (Betts, 2010; Raudenbush & Byrk, 1988). Because multilevel linear modeling allows individual participants to retain their own intercepts and growth estimates, one is able to examine within-person change over time in addition to between-person differences in change. Singer and Willett (2003) also noted the appropriateness of multilevel linear modeling in school research due to its ease of accommodating missing data (e.g., absences, school cancellations).

Theoretical Conceptualizations of Generality of Behavior

The importance of assessing generality of behavioral change as a result of intervention has been well established in the academic literature for decades (e.g., Baer, Wolf, & Risley, 1968; Fox & McEvoy, 1993). Behavioral changes are considered to have generality if, under non-training conditions, they (a) are durable over time, (b) appear in a wide variety of environments, or (c) extend to other related, non-trained responses (Baer et al., 1968; Stokes & Baer, 1977). On the contrary, when behaviors similar to intervention effects are produced only under conditions similar to those of the intervention itself, generality of behavior cannot be claimed (Stokes & Baer, 1977). Although “generality of behavior” and “generalization” are often used synonymously (Cooper, Heron, & Heward, 1987), it is important to distinguish between the

terms so as not to suggest any desirable behavior change that occurs in a non-training setting is the result of a single process. Generality of behavior change, then, can be conceptualized as consisting of multiple components (i.e., generalization and maintenance), which are discussed in detail below.

Traditionally, maintenance and generalization are conceptualized as separate, hierarchical processes. For example, Ardoin (2006) asserted that the generalization of a skill is likely dependent on the extent to which that knowledge is first maintained. This theoretical viewpoint aligns with the instructional hierarchy model of learning set forth by Haring, Lovitt, Eaton, and Hanson (1978). According to this model, students progress though a learning hierarchy when learning a new skill. During the acquisition stage, when new behaviors are not part of a student's repertoire, teachers must focus on increasing the occurrence of accurate responses, and decreasing the occurrence of inaccurate responses. Once students respond accurately the vast majority of the time, they enter the fluency stage in which they respond quickly and accurately. From this point, Haring and colleagues (1978) proposed that teaching efforts should focus on promoting generalization across multiple stimuli (e.g., settings). The final stage of the instructional hierarchy assumes that instructional efforts should focus on teaching students to adapt their skills to new situations.

Similar to Haring and colleagues' (1978) conceptualization of generality, Rivera and Bryant (1992) viewed maintenance and generalization as separate, hierarchical stages in the learning process, as the ability to maintain skills was assumed to precede the ability to generalize skills. In the context of education, they argued that the learning process begins with the acquisition stage, in which students know little about the specified skill. Through practice opportunities with teacher guidance, students are thought to transition to the proficiency stage, in

which they understand how to do the skill, but need to practice using the skill to develop competency. Students reach the maintenance stage of learning when they are able to independently practice mastered skills. Once students have mastered the maintenance stage, they are thought to enter the generalization stage of learning in which they are able to apply their acquired skills across different people, materials, and settings. Finally, once generalization is established, students are thought to enter the adaptation stage when they can adapt their knowledge to problem-solve.

Generalization. Generalization consists of two behavioral processes: (a) stimulus generalization and (b) response generalization. According to Cooper et al. (1987), stimulus generalization occurs when a learner's performance of a target behavior improves in environments that differ from the original training environment. That is, stimulus generalization refers to generality across settings, people, and conditions. As a general principle, the more a given stimulus resembles the training stimuli, the greater the likelihood that stimulus generalization will occur. For example, students will likely demonstrate greater evidence of stimulus generalization of writing fluency behaviors on writing tasks that are similar to writing tasks on which they were trained. Conversely, stimulus generalization is less likely to occur when given stimuli configurations differ significantly from stimuli that were present under training or intervention conditions.

Response generalization is a behavioral process in which a variety of functional responses, which have not been reinforced in a training condition, are emitted in addition to the trained responses (Cooper et al., 1987). The occurrence of response generalization implies that a given stimulus, to which an individual was trained to produce a particular response, evokes similar but different responses. For example, Campbell, Brady, and Linehan (1991) assessed for

response generalization by examining students' ability to transfer skills of identifying words that needed to be capitalized to correctly capitalizing words in their own writing compositions.

Within the past decade, the American Psychological Association's Division 16 and the Society for the Study of School Psychology created a task force to develop a knowledge base regarding evidence-based interventions that have direct relevance to the field of school psychology (see Kratochwill & Stoiber, 2002). It was concluded that providing evidence of generalization of treatment effects is a key component that should be incorporated into treatment studies. Specifically, it was determined that treatment studies should provide evidence of the extent to which participants are able to generalize their newly acquired skills across settings and across persons.

Maintenance. Maintenance can be conceptualized as the extent to which an individual's desired behavior change, after beginning to be emitted in non-training settings, persists after all or some of the intervention has ceased (Cooper et al., 1987). The importance of assessing whether treatment has a lasting effect on behavior (i.e., maintenance) has been amply established in the scientific literature. In particular, a number of early works (e.g., Baer, et al. 1968; Lovitt, 1975) highlight the pragmatic utility of examining maintenance. Specifically, Baer et al. (1968) claimed that a crucial aspect of applied research is the evaluation of a behavior's generality (e.g., the durability of a behavior over time). Similarly, Lovitt (1975) criticized research that failed to examine the retention of behavioral change, and suggested that researchers routinely report retention data.

Along with generalization, the task force on evidence-based interventions emphasized the importance of assessing for durability of intervention effects (Kratochwill & Stoiber, 2002). Specifically, the authors determined that strong evidence of maintenance assessment is

established in studies that (a) conduct follow-up assessments over multiple intervals (e.g., 6 months and 1 year), (b) conduct follow-up assessments with all participants from the original sample, and (c) use measures that are similar to those used to analyze data from the original intervention. Promising evidence is achieved in studies that (a) administer follow-up assessment at least once (e.g., 6 months), (b) conduct follow-up assessments with a majority of the participants included in the original sample, and (c) use measures that are similar to those used to analyze data from the original intervention. Finally, weak evidence of maintenance assessment is derived from studies that (a) conduct follow-up assessments at least once (e.g., 6 months) and (b) conduct follow-up assessments with some participants from the original sample.

In addition to satisfying recommendations made by the American Psychological Association, measuring generalization and maintenance provides evidence of construct validity associated with both the intervention and the assessment (D'Agostino, 2005). Specifically, strong relationships between measurements made during intervention sessions and generality sessions would serve as evidence supporting the notion that a single construct was measured.

Programming and Assessing Generality of Behavioral Change

In 1977, Stokes and Baer reported techniques to program and assess generality of behavior change based on their review of the extant scientific literature in that area. The reviewed research was categorized by the techniques used to assess or program generality. Specifically, these researchers found that literature of that time fell into nine categories: (a) train without programming generality and hope for positive results, (b) program generality in the aftermath of an intervention when no evidence of generality was found, (c) incorporate natural maintaining contingencies into the intervention procedures, (d) train sufficient exemplars, (e) loosen stringent control over training methods, (f) use indiscriminable contingencies, (g)

program common stimuli into training and generality sessions, (h) mediate generality, and (i) reinforce instances of generality. A majority of literature was found to fall into the "Train and Hope" category, suggesting that at that time most studies examining generality of behavior change did not specifically attempt to program generality; rather, researchers simply trained a behavior to acquisition, then assessed for generality of behavior change.

Stokes and Osnes (1989) sought to refine the work by Stokes and Baer (1977). Rather than reviewing research that simply assessed for generality in the aftermath of an intervention, Stokes and Osnes focused on literature that suggested promising techniques of programming generality of behavior change to maximize the likelihood of its occurrence. Tactics of generality programming were categorized into three overarching principles: (a) exploit current functional contingencies, (b) train diversely, and (c) incorporate functional mediators. Each category consisted of four tactics that were suggested for use in the programming of generality.

Research studies that were categorized under Stokes and Osnes' (1989) principle of exploiting current functional contingencies focused on the manipulation of a behavior's antecedents and consequences to promote generality of behavior change. The first tactic proposed under this category was to train behaviors that are likely to contact powerful reinforcers in the natural environment. Thus, it was suggested to incorporate reinforcers that are inartificial and naturally occurring into the intervention program. In the event that reinforcers are not prevalent enough in the natural environment to promote generality of behavior change, it was suggested to teach participants to recruit natural consequences, thereby becoming active agents of their own behavior change. For instance, Stokes, Fowler, and Baer (1978) taught children with behavior disorders to increase the quality of their academic performance, and then occasionally ask their teachers questions such as, "How is this?" This cued their teachers to notice and

reinforce their behavior. A third technique suggested in this category was to attempt to decrease the frequency of maladaptive behavior by extinguishing its reinforcing consequences so that more appropriate behaviors can be developed and maintained. A final technique to program generality of behavior change described under this principle was to reinforce instances of generality. Importantly, Campbell and Willis (1978) demonstrated the effectiveness of this technique on fifth-grade students' essay writing. By using social and token reinforcement for occurrences of generalization (i.e., variability from trained behaviors), these writing behaviors were shown to increase and maintain over time.

The second general principle of generality programming found in the literature by Stokes and Osnes (1989) was training diversely. This category supported the notion of using less stringently controlled methods during training conditions to program generality of behavior change. Specifically, this category focused on allowing variations in antecedent stimuli, responses, and consequences during training conditions. This principle likely promotes generality because "focused training frequently has focused effects" (Stokes & Osnes, 1989, p. 344), and thus, making training conditions less focused may aid in the production of behavioral generality. One tactic proposed in this category was to use a sufficient number of stimuli conditions during training (i.e., use sufficient stimulus exemplars). For example, using multiple trainers and training in multiple settings was noted to result in greater generality of behavior change. Stokes and Baer (1977) found that a small number of stimulus exemplars (i.e., two trainers or two settings) was frequently sufficient to aid generality. Stokes and Osnes also recommended incorporating into training procedures a subset of responses from the particular class that is to be generalized and maintained (i.e., use sufficient response exemplars). A third tactic of training described by Stokes and Osnes involved making antecedents less discriminable

so the participant does not perform the desired behavior only under a particular set of circumstances. Similarly, Stokes and Osnes recommended making consequences less discriminable by (a) using intermittent reinforcement schedules, which are particularly useful for facilitating maintenance, (b) delaying the presentation of consequences, or (c) reducing the predictability of the intervention agent's presence to deliver consequences.

A final principle of programming generality of behavior change described by Stokes and Osnes (1989) was incorporating functional mediators. That is, incorporating some type of common stimulus – possibly a discriminative stimulus – in both the training setting and the generality setting. Incorporating common salient physical stimuli (e.g., materials, reinforcers) and social stimuli (e.g., training agent, peers) into the training and generality conditions have been shown to increase the occurrence of generality of behavior change. Similarly, incorporating salient self-mediated physical stimuli (e.g., a notebook or lecture notes indicating how to perform under certain conditions) and verbal and covert stimuli (e.g., self-instructions, goal-setting) into both the training and generality settings has been shown to improve generality outcomes (Guevremont, Osnes, & Stokes, 1988; Kelley & Stokes, 1984).

Research Examining Generalization of Writing Skills

A number of research studies have examined the generalized effects of writing interventions among elementary-aged populations. However, due to the variability in interventions and outcome measures reported in these studies, it is difficult to make definitive conclusions regarding this area of research. Specifically, of 36 writing intervention studies that examined generalization of writing skills, 13 different types of interventions were examined, and more than 20 different dependent measures were reported. Below is a discussion of the extant scientific literature examining the generalized effects of different types of writing interventions.

This briefing begins with a synopsis of the most frequently measured writing outcomes, and is followed by a discussion of different types of interventions that have been examined.

Writing intervention outcome measures. A variety of writing outcome were measured in the literature examining generalization of writing intervention effects. Among the most frequently measured were composition quality ratings (e.g., Graham & Harris, 1989; Graham & MacArthur, 1988; Medcalf, Glynn, & Moore, 2004; Monroe & Troia, 2006; Regan, Mastropieri, & Scruggs, 2005), capitalization (e.g., Campbell et al., 1991; Hermann, Semb, & Hopkins, 1976), spelling (e.g., Diaz, McLaughlin, & Williams, 1990; Hanna, de Souza, de Rose, & Fonseca, 2004; Pratt-Struthers, Struthers, & Williams, 1983), and handwriting (i.e., Robin, Armel, & O'Leary, 1975; Trap, Milner-Davis, Joseph, & Cooper, 1978). Although composition length was also a commonly measured outcome (e.g., Graham & Harris, 1989; Graham & MacArthur, 1988; Medcalf et al., 2004, Regan et al., 2005, Van Houten, 1979), in some studies it was not necessarily an indicator of fluency. Specifically, only two of these studies (Medcalf et al., 2004; Van Houten, 1979) controlled the amount of time for which students wrote, and were thus able to examine the rate of the total number of words written.

Generalization and Self-Regulated Strategy Development. Notably, a majority of research examining generalization of writing skills was conducted in the area of Self-Regulated Strategy Development (SRSD; e.g., Graham & Harris, 1993; Graham, Harris, & Mason, 2005; Harris & Graham, 1996), an intervention designed to promote academic strategic behavior, knowledge, and motivation. A variety of instructional procedures designed to promote maintenance and generalization were incorporated into the SRSD model. Specifically, the SRSD instructional model includes steps to (a) ensure the purpose of each target strategy is clear to the child, (b) teach strategies to the point of fluency to increase the likelihood of correctly applying

the strategies to different types of writing prompts, (c) facilitate strategy use through self-talk, (d) promote self-monitoring and evaluation for feedback purposes, (e) pair performance gains with strategy use, (f) encourage maintenance and generalization of skills, and (g) discuss when, where, and how to use strategies (Harris & Graham, 1996). Graham (2006) conducted a meta-analysis of 39 studies that examined SRSD. Of the 19 studies that reported generalization data, 18 found evidence of successful generalization of a variety of writing behaviors (e.g., total number of words written, revision skills, composition quality) following instruction of this strategy. For studies that employed a group design, effect size was calculated using Glass's Δ (i.e., the posttest mean of the control group was subtracted from that of experimental group, and this was divided by the standard deviation of the control group). For studies that employed a single-case design, effect size was calculated by computing the percentage of non-overlapping data points. The percentage of non-overlapping data was defined as the proportion of data points in the treatment condition that exceeded the most extreme data point in the baseline condition. Generalization effect sizes ranged from .20 to 3.34 for studies utilizing group design, and the percent of non-overlapping data for single case design studies ranged from 33% to 100%.

Generalization and performance feedback interventions. Four studies have examined the generalized effects of writing skills following performance feedback interventions; however, only one examined performance feedback in isolation. Trap et al. (1978) examined the effects of verbal and visual feedback on the correct cursive letter strokes of 12 first-grade students. Although the percentage of correct letter strokes increased following intervention, results did not show a consistent pattern of generalization to untrained and unpracticed letter strokes. In contrast, students who participated in three studies (Sajwaj & Risley, 1970; Schunk & Swartz, 1993; Van Houten, 1979) that examined performance feedback in combination with other

intervention components (e.g., goal-setting, explicit timing) demonstrated evidence of generalization of writing behaviors.

Generalization and self-instructional training interventions. Self-instructional training involves teaching students composition strategies that can later be used when writing independently. For example, Graham and MacArthur (1988) taught students to produce longer, better quality compositions by training and modeling self-instructional strategies (e.g., re-reading the essay, ensuring the message is clear, asking oneself if it makes sense, questioning whether any more material can be added). Evidence of generalization effects of this type of intervention has been inconsistent. When examining this instructional strategy with three 10- to 11-year-old students with learning disabilities, Graham and MacArthur found that students' ability to write longer, higher quality essays on a word processor generalized to writing essays with pen and paper. Similarly, Graham and Harris (1989) found that this type of training resulted in the generalization of three sixth-grade students' improvement in writing skills (i.e., writing essays with more total words, more functional elements and grammar elements, better quality, and more coherence) to a new setting, teacher, and genre. Conversely, Robin et al. (1975) did not find evidence of generalization of printing (i.e., handwriting) ability in 20 kindergarten students following self-instructional training.

Generalization and peer tutoring interventions. Two studies (Campbell et al., 1991; Medcalf et al., 2004) examined elementary-aged students' ability to generalize writing behaviors following peer tutoring. Findings from both studies indicated evidence of generalization of writing skills as a result of peer tutoring. Specifically, Campbell et al. (1991) found evidence of stimulus and response generalization of correct capitalization ability in three second-grade students with mental retardation and learning disabilities. Medcalf et al. (2004) examined the

compositions of seven low-achieving first-grade students, and found evidence of stimulus generalization in terms of increased rate, accuracy, and message clarity.

Research Examining Maintenance of Writing Skills

Similar to the scientific literature base examining the generalization of writing intervention effects, studies examining elementary-aged students' ability to maintain writing gains as a result of intervention have assessed the impact of a variety of instructional techniques on a variety of writing outcome measures. Of 33 reviewed studies, approximately 10 different types of interventions were examined, and more than 15 writing outcomes were measured. Below is a discussion of the existing literature examining maintenance of writing intervention effects. Commonly measured writing outcomes of these studies will first be discussed, followed by a review of the writing interventions that have been most frequently examined in this area.

Writing intervention outcome measures. Among studies that examined maintenance of writing intervention effects, a wide array of outcome variables was measured. The most commonly measured writing outcomes included spelling (e.g., Diaz et al., 1990; McGuffin, Martz, & Heron, 1997; Moore, Heward, & Alber, 1998), composition quality (e.g., Graham & Harris; 1989; Graham & MacArthur, 1988; Medcalf et al., 2004; Regan et al., 2005), and composition length (e.g., Graham & Harris; 1989; Graham & MacArthur, 1988; Medcalf et al., 2004; Regan et al., 2005; Van Houten, 1979). Again, only two of the studies that examined composition length (Medcalf et al., 2004; Van Houten, 1979) truly measured writing fluency (i.e., rate of total words written per a specified amount of time). Other studies that examined length of students' compositions did not control for time allocated to the writing task (e.g., Graham & Harris, 1989) or simply estimated length based on visual inspection of the composition (i.e., Regan et al., 2005).

Maintenance and SRSD. Consistent with those studies examining generalization of writing intervention effects, the scientific literature base assessing maintenance of such effects has been dominated by studies investigating the impact of SRSD. In a meta-analysis of these studies, Graham (2006) found that of the 20 studies that examined maintenance of writing gains, 18 reported evidence of successful maintenance. In the reviewed studies, maintenance data were collected between 2 and 15 weeks. Effect sizes were calculated based on computational procedures identical to those described previously (see Generalization and Self-Regulated Strategy Development). For studies utilizing a group design, effect sizes for generalization outcomes ranged from .35 to 4.82, and for studies utilizing a single case design, the percent of non-overlapping data ranged from 33% to 100%.

Maintenance and self-instructional training interventions. Two studies (Graham & Harris, 1989; Graham & MacArthur, 1988) have examined students' ability to maintain writing skills that were developed as a result of self-instructional training. Both of these studies, which were described above, reported evidence of maintained treatment effects. Specifically, Graham and MacArthur found that students were able to maintain their ability of writing longer, higher quality essays for up to 9 weeks. Additionally, Graham and Harris (1989) found that students' ability to write coherent, higher quality, and longer essays that incorporated more functional and grammar elements was maintained over periods of 6, 11, and 12 weeks.

Maintenance and self-correct interventions. Two studies (McGuffin et al., 1997; Moore et al., 1998) examined students' ability to maintain spelling gains following the implementation of a self-correct intervention. Specifically, students were able to preview a list of spelling words, and then were instructed to write the spelling words upon auditory presentation from a cassette tape. Following this task, students were again presented with the visual list of

spelling words, and were instructed to self-correct their answers. McGuffin et al. (1997) found that at intervals of 6, 8, and 10 weeks following this procedure, five out of six third-grade students at risk for spelling failure maintained spelling gains. Similarly, Moore et al. (1998) found that four out of five students with learning disabilities aged 11 to 12 years were able to demonstrate maintenance of spelling gains on probes administered 1 week following intervention.

Maintenance and performance feedback interventions. Two studies (Schunk & Swartz, 1993; Van Houten, 1979) examined maintenance of writing skills following the implementation of performance feedback interventions. Schunk and Swartz (1993) reported evidence of maintenance of writing skills (i.e., paragraph quality) for a group of 40 fourth-grade students following an intervention that integrated performance feedback, strategy instruction, and goal setting. Van Houten (1979) also found evidence of maintenance of the number of words written per minute following an intervention that combined performance feedback and explicit timing. Although both studies found evidence of maintenance, similar to a limitation of the generalization literature, neither of these studies used methods that allowed for the assessment of performance feedback in isolation.

Summary of Research Examining Generalization and Maintenance of Writing Skills

Similar to Stokes and Baer's (1977) findings, the majority of studies examining generality (i.e., generalization and maintenance) of writing intervention effects reported evidence of successful generality. As Stokes and Baer proposed, this plethora of encouraging data suggests (a) strategies that support the occurrence of generality were incorporated into training or (b) a possible tendency to underreport generality data if no evidence of generality is found.

Because all studies incorporated into training at least one generality programming tactic (Stokes & Osnes, 1989), it appears the former possibility may be more explicative in this case.

Although a variety of research studies have examined elementary-aged students' ability to generalize and maintain skills developed as a result of writing interventions, this area of scientific literature is limited for a number of reasons. First, only one intervention (i.e., SRSD) has received a considerable amount of attention in terms of replication of generality effects. Specifically, 25 studies have examined generality effects of SRSD. Second, very few studies (18%) have examined typically developing, general education students' ability to generalize and maintain writing skills. Third, most studies have relied on single-case design (71%), which, due to limited systematic replications, limits the external validity. Finally, few studies (9%) have examined this in the context of a performance feedback intervention, and only one has studied performance feedback in isolation. Importantly, no studies have examined the generality of students' writing fluency skills following an isolated performance feedback intervention. This is disconcerting given the recent efforts by professional organizations (i.e., Division 16 and the Society for the Study of School Psychology) to emphasize the importance of including generalization and maintenance data in studies examining intervention effects.

Purpose of the Current Study

The goal of the proposed study was to extend existing research in the area of performance feedback intervention effects on elementary-aged students' writing fluency by examining whether the intervention investigated by Eckert et al. (2006) resulted in generality of behavior change. A primary aim of this study was to examine whether students who received a performance feedback intervention demonstrated greater writing fluency improvement than students assigned to a practice-only control condition. Changes in students' writing fluency (i.e.,

growth trajectories) were compared across the two conditions. Because previous research has shown that both practice and performance feedback can increase students' writing fluency skills (Eckert et al., 2006; Harris et al., 1994), it was hypothesized that students in both conditions would demonstrate improvements over time. However, similar to Eckert et al.'s (2006) results, it was expected that students assigned to the performance feedback intervention would make significantly greater gains than students assigned to the practice-only condition.

In an attempt to examine whether the performance feedback intervention examined by Eckert et al. (2006) resulted in generality of behavior change, a central and primary aim of this study was to examine whether students who received performance feedback demonstrated greater writing fluency gains on (a) a stimulus generalization measure, (b) a stimulus-response generalization measure (i.e., a task that incorporated aspects of both stimulus generalization and response generalization), and (c) maintenance measures than students assigned to the practice-only control condition. Because the performance feedback intervention incorporated a number of tactics that have been shown to program generality (i.e., the use of sufficient stimulus exemplars in the form of multiple research assistants, making consequences less discriminable by delaying presentation, incorporating common salient physical and social stimuli into both the training and generality settings; Stokes & Osnes, 1989), it was hypothesized that students in the performance feedback condition would demonstrate evidence of generalization and maintenance on measures designed to assess these areas to a greater extent than students assigned to the practice-only control condition. Specifically, it was hypothesized that (a) when controlling for baseline performance on a stimulus generalization measure, students assigned to the performance feedback condition would perform significantly better than students assigned to the practice-only control condition on a post-intervention stimulus generalization measure, (b) when controlling

for baseline performance on a stimulus-response generalization measure, students assigned to the performance feedback condition would perform significantly better than students assigned to the practice-only control condition on a post-intervention stimulus-response generalization measure, and (c) the correlations between student performance on the final intervention probe and three maintenance probes would be higher for those assigned to the performance feedback condition than for those assigned to the practice-only control condition.

A secondary aim of this study was to examine shifts in students' instructional level as a result of receiving performance feedback. Research suggests that both practice and performance feedback have a positive impact on the instructional level (i.e., frustrational, instructional, mastery; Shapiro, 2004) at which students write (Eckert et al., 2006). Thus, it was hypothesized that students in both the performance feedback and practice-only condition would advance in their writing instructional level (e.g., students writing at the frustrational level would experience fluency gains that advanced them to the instructional level). In line with Eckert et al.'s (2006) study, students receiving a combination of practice and performance feedback (i.e., students in the performance feedback condition) were expected to make greater shifts in instructional level than students in the practice-only condition.

Another secondary aim of this study was to examine the ability of intervention factors (i.e., instructional level at baseline session, instructional level at final intervention session, and writing fluency growth over time) and student demographic factors (i.e., absenteeism, ethnicity, and acceptability of intervention score) to predict generalization and maintenance of intervention effects. Because the acquisition and fluency of a skill have been theorized to precede one's ability to demonstrate generality of that skill (Haring et al., 1978), it was hypothesized that instructional level and writing fluency growth over time (i.e., indicators of acquisition and

fluency) would significantly predict students' ability to generalize and maintain intervention effects over time. Furthermore, it was hypothesized that students' absenteeism, ethnicity, and acceptability of the intervention would significantly predict generalization and maintenance.

Method

Participants and Setting

Upon Institutional Review Board approval from Syracuse University and the participating school district, third-grade students enrolled in general education classrooms were invited to participate in the study. Parental permission (Appendix A) was obtained prior to student participation in the study and student assent (Appendix B) was sought. Students for whom consent and assent were obtained were then screened for eligibility and invited to participate in the study. The eligibility criteria included: (a) not experiencing severe motor deficits that precluded students from composing written stories; (b) not experiencing severe cognitive deficits that resulted in eligibility for special education services; (c) English was the primary language spoken by the child; (d) not classified as Learning Disabled; (e) not having a one-to-one instructional aide or a Section 504 plan indicating additional instructional modification; (f) demonstrating minimum proficiency by writing at least seven words on a baseline measure (described in Measures); and (g) legibly scribing a subset of letters from the alphabet. Ineligible students and those students without consent participated in an alternate instructional activity identified by their teacher.

To ensure adequate power in testing the differences in growth trajectories (i.e., slope) between the two conditions over seven sessions of data collection, an a priori power analysis was conducted based on procedures developed by Diggle, Liang, and Zeger (1994) for multi-level modeling. Sample size was calculated by setting α equal to 0.05 and power equal to 0.80. Based

on pilot work by Eckert et al. (2006), the sample size was calculated to detect a minimum meaningful difference in slopes of 0.60. The results indicated that 32 third-grade participants per condition must be included, resulting in an overall required sample size of at least 64 participants. A total of 103 third-grade participants were included in this study (51 in the experimental condition and 52 in the control condition), which exceeded the requirements set by the power analysis.

All students were recruited from two public elementary schools in a large school district that was located in a moderately sized urban city in the northeast. The schools were selected due to proximity to the university, and the sample of students represented a sample of convenience. All sessions took place in the students' general education classrooms during a 30-minute block of time identified by the classroom teachers. Detailed information regarding the recruitment, enrollment, and intervention allocation are reported according to the Consolidated Standards of Reporting Trials guidelines (Figure 3; Moher, Schulz, & Altman, 2001). Student demographic variables and school characteristics were also examined, and they are reported and included in subsequent analyses.

According to the most recent New York State School Report Card, which published demographic data for the 2009-2010 school year, 392 kindergarten through fifth-grade students were enrolled in the first school. Of the 392 students enrolled in this school, 88% were eligible for free or reduced-priced lunch. The majority of students enrolled in this school were identified as Black or African American (63%). A smaller percentage were identified as White (22%), Hispanic or Latino (13%), American Indian or Alaska Native (2%), and Asian or Native Hawaiian/Other Pacific Islander (1%).

The second school was comprised of 831 students who were enrolled in kindergarten through eighth-grade. A total of 65% of the students enrolled in this school were eligible to receive free or reduced-priced lunch. In terms of race and ethnic composition, the majority of students were identified as either White (55%) or Black or African American (36%). A smaller percentage were identified as Hispanic or Latino (6%), American Indian or Alaska Native (2%), and Asian or Native Hawaiian/Other Pacific Islander (1%).

All participating teachers (N = 6) held at least a bachelor's degree, and five (83%) held master's degrees in education. Two teachers (33%) held an additional certification in literacy, and two teachers (33%) held degrees in special education. The mean number of years of teaching experience was 17.6 (range, 3 to 34).

Experimenters

School psychology doctoral students and advanced undergraduate psychology majors served as research assistants. Prior to data collection, all research assistants were required to complete a formal training in research ethics, as required by Syracuse University. This training (i.e., Collaborative Institute Training Initiative) provides online basic courses in the protection of human research subjects. All research assistants were required to submit documentation that they successfully passed the Social and Behavioral Focus and Responsible Conduct of Research courses. This documentation was submitted to the Institutional Review Board.

All research assistants received training in administering dependent measures, scoring dependent measures, conducting procedural integrity observations, and completing data entry. In addition, research assistants were provided with procedural scripts for administering dependent measures, a manual detailing the scoring procedures for the dependent measures, and procedural scripts for conducting procedural integrity observations. They received training on all

procedures, followed by opportunities to practice and receive feedback on scoring writing probes. All research assistants were required to demonstrate 100% proficiency administering dependent measures, scoring dependent measures, and conducting procedural integrity observations.

Materials

Several measures were administered to evaluate participants' skills in written expression and writing fluency. Writing fluency was primarily examined with Curriculum-Based Measurement probes in Written Expression during intervention, maintenance, and stimulus generalization sessions. A measure similar to students' typical writing classwork was used to assess the combination of stimulus and response generalization. Secondary measures, specifically, the paragraph-writing portion of the Test of Written Language-Third Edition (Hammill & Larsen, 1996), a paragraph copying task from the Monroe-Sherman Group Diagnostic Reading Aptitude and Achievement Test (Monroe & Sherman, 1966), and an informal measure of handwriting were used to assess students' global writing abilities. A student intervention acceptability measure and a teacher questionnaire were administered for use in exploratory analyses and for descriptive purposes.

Curriculum-Based Measurement probes in Written Expression. Curriculum-Based Measurement (CBM; Deno, 1985) is an assessment tool in which brief measures of academic behavior are administered repeatedly to examine skill development over time. This measurement tool can assess students' skills in a number of curricular areas. For the purposes of this study, CBM in Written Expression was used as a measure of writing fluency.

CBM probes in Written Expression (CBM-WE) were developed based on procedures outlined by Shapiro (2004). To assess writing fluency with CBM, students are provided with one

probe containing a story starter (e.g., “I was talking to my friends when all of a sudden ...”) and are instructed to spend 1 minute planning a story based on the story starter. Students are then instructed to spend 3 minutes writing a narrative story, and are prompted by the assessor to continue writing for the entire 3 minutes.

A total of 12 probes were used over the course of this study (i.e., one probe during each intervention session, two probes for pre- and post-stimulus generalization measures, and one probe during each maintenance session). Each probe contained a story starter that had been previously evaluated for use with elementary-aged students (AIMSweb®, 2004; McMaster & Campbell, 2006). The story starters were each comprised of a short sentence fragment, and were intended to provide students with an idea for writing a narrative story. As is further described below, the 10 intervention and maintenance CBM probes consisted only of the pronoun “I” (i.e., self-referenced; e.g., “I found a note under my pillow that said…”). The two stimulus generalization CBM probes consisted only of the pronouns “she” and “he” (i.e., others-referenced; e.g., “As he opened the door the…”). A complete listing of story starters that were used in this study is provided in Table 1.

During each intervention session, one probe was presented in the form of a writing packet to each student. The first page of the packet contained the students’ identifying information (Appendix C). In an attempt to reduce the likelihood of student previewing of the story starter, the second page of the packet contained a picture of a stop sign in the middle of the page (Appendix D). The third page of the packet contained individualized performance feedback sheets (described below). The remaining pages of the packet contained the CBM-WE probe materials. The probe materials included: (a) one page containing a story starter written across the

top of the page and a stop sign at the bottom (Appendix E), (b) one page containing the story starter with compositional lines (Appendix F), and (c) one page containing compositional lines.

Although numerous CBM-WE outcome measures have been evaluated as possible indicators of writing fluency, total words written, words spelled correctly, and correct writing sequences are the three that are the most commonly and appropriately used to assess writing fluency among elementary-aged children (Espin et al., 2000). Two sources (McMaster & Espin, 2007; Powell-Smith & Shinn, 2004) provide comprehensive reviews of studies that explored the technical adequacy of total words written, words spelled correctly, and correct writing sequences. The resulting reliability and validity coefficients are summarized in Table 2. Overall, reliability coefficients (range, r = .51 to .99), as well as interscorer agreement (range, 91% to 99%) for total words written and correctly spelled words were moderate to high. Validity studies indicated that correct writing sequences were more highly correlated with criterion measures (e.g., holistic and informal teacher ratings, Test of Written Language [Hammill & Larsen, 1996], Minnesota Basic Skills Test [Minnesota Department of Children, Families, and Learning & NCS Pearson, 2002]) than either total words written or words spelled correctly (range, r = 0.18 to 0.85). Similarly, in comparison to the other two metrics, correct writing sequences has been shown to be the most accurate measure when monitoring student growth over time (Hubbard, 1996).

Stimulus-response generalization probe. Students were administered a writing task that closely resembled their typical classwork in the subject of writing, as identified by their classroom teachers (see Appendix G). The exact nature of this teacher-administered, compare-and-contrast writing task was determined upon contact with classroom teachers prior to beginning the intervention. Specifically, all six classroom teachers identified the Treasures

(Macmillan McGraw-Hill, 2006) compare-and-contrast test as the end-of-year exam that all third-grade students would be required to complete, and as such the teachers reported frequently administering past exams as classroom writing assignments. To examine the extent to which writing fluency gains transferred to this writing assignment, the classroom teachers administered a modified version of the Treasures compare-and-contrast test. Using a procedural script (Appendix H), the teachers visually presented the students with a topic. Students were given 3 minutes to plan their composition and 10 minutes to write their compare-and-contrast essay. This measure differed from the Treasures compare-and-contrast test in that (a) it was a timed test and (b) students were not allowed to begin writing their composition until the planning period had ended.

Test of Written Language – Third Edition. The Test of Written Language – Third Edition (TOWL-3; Hammill & Larsen, 1996) is a standardized, norm-referenced test designed to measure writing abilities of children aged 7 to 17. The TOWL-3 is comprised of eight subtests designed to measure writing competence using both spontaneous and contrived formats. Contrived format subtests measure isolated aspects of writing, such as spelling, punctuation, and capitalization. Spontaneous format subtests measure a variety of writing skills (e.g., spelling punctuation, format, plot, sentence structure, readability) within a meaningful task. For this format, students are instructed to write a story about a picture that has been provided.

In this study, the Spontaneous Writing subtest of the TOWL-3 (Appendix I) was administered to examine students' general writing abilities. This subtest required students to look at a picture, then plan and write a story based on the picture for 15 minutes. The story was then evaluated in the areas of (a) Contextual Conventions (e.g., punctuation, capitalization, and

spelling), (b) Contextual Language (e.g., vocabulary, grammar, and sophistication of sentence construction), and (c) Story Construction (e.g., plot, character development, and style).

The technical adequacy of the TOWL-3 has been primarily evaluated by the test developers and the psychometric properties have been reported in the test manual. The internal consistency of the Spontaneous Writing Quotient for eight- and nine-year-old children is high (r = .90). Additionally, interscorer and alternate-form reliability is high (i.e., coefficients are greater than .80). A comparison of scores from the Spontaneous Writing Composite and another standardized measure of student performance (i.e., Comprehensive Scales of Student Abilities) suggests a moderate association in terms of criterion-related validity (Hammill & Hresko, 1994; r = .50). Burns and Symington (2003) examined the criterion-related validity by correlating Spontaneous Writing Quotient scores with teacher ratings of writing achievement. Results indicated the criterion-related validity for the Spontaneous Writing Quotient was low (range, r = .39 to .48).

Paragraph copying task from the Monroe-Sherman Group Diagnostic Reading Aptitude and Achievement Test. The paragraph copying task from the Monroe-Sherman Group Diagnostic Reading Aptitude and Achievement Test (Monroe & Sherman, 1966) was administered at baseline (see Appendix J) to provide an initial indicator of orthographic skill. Normative data are based on the number of words copied accurately. This task is the only paragraph copying task with published normative data for elementary-aged children. Psychometric properties and published norming procedures of this measure are not available. However, this measure has been shown to be a significant predictor of overall writing ability and writing fluency (Berninger, Hart, Abbott, & Karovsky, 1992; Graham et al., 1997).

Informal measure of handwriting. To determine the legibility of students' handwriting, participants will be asked to print a set of 10 lowercase letters from the alphabet (i.e., f, c, r, m, v, y, i, h, e, o). These 10 letters were randomly chosen utilizing a random numbers generator after excluding the commonly reversed letters 'b' and 'd.' The measure was developed by the author and no psychometric evidence is currently available (see Appendix K).

Student intervention acceptability measure. A brief acceptability measure was administered to the students assigned to the performance feedback condition to assess their perceptions of the intervention used in the study. The purpose of this measure was to obtain a short, descriptive evaluation of the students' perceptions of the specific procedures used during the study. Although this measure was developed in lieu of a previously published survey, these particular questions were based on the Children's Intervention Rating Profile (Witt & Elliott, 1983) and have been used in previous studies examining the acceptability of academic interventions.

This assessment included a series of questions using a 5-point Likert-type response system. Response options ranged from "*not at all*" to "*very, very much*." Students assigned to the performance feedback group received a five-page packet (Appendix L) containing eight questions regarding their attitudes towards writing, their perceptions of procedures used in the intervention, and their perceptions toward receiving feedback. The first four and last two questions were also administered to the students assigned to the practice-only control condition. Because all of the questions were developed by the author for use in this study, Cronbach's alpha reliability coefficients were calculated. The measure was found to have adequate internal consistency for the entire sample (6 items, $\alpha = .70$), the performance feedback condition (8 items, $\alpha = .77$), and the practice-only condition (6 items, $\alpha = .71$).

Intervention Rating Profile-15. For descriptive analysis purposes, all participating classroom teachers were asked to complete an adapted version of the Intervention Rating Profile-15 for Teachers (IRP-15; Martens, Witt, Elliott, & Darveaux, 1985; Appendix M). The abridged version of this rating scale consists of 15 questions that are rated on a 6-point Likert-type scale, with higher scores indicating higher treatment acceptability. For the purpose of this study, the words "problem behavior" were modified to read "writing difficulties" on the questionnaire. Because this measure was adapted for the purposes of this study, Cronbach's alpha reliability coefficients were calculated. Results indicated that this measure of acceptability demonstrated excellent internal consistency (15 items, $\alpha = .95$).

Teacher questionnaire regarding writing instruction. Participants' classroom teachers were asked to complete the Writing Orientation Scale (Graham, Harris, MacArthur, & Fink, 2002; see Appendix N) for descriptive analysis purposes only. This scale was designed to measure teachers' beliefs about writing instruction. There is some evidence indicating that teachers' beliefs greatly influence their practices and their students' outcomes (Graham et al., 2002). Writing instruction across classrooms is highly variable, so a description of the writing instruction provided in the participants' classrooms were measured and reported. The items on the Writing Orientation Scale load onto three factors: (a) correct writing, which accounts for 15% of the total variance in the scale, (b) explicit instruction, which accounts for 12% of the total variance, and (c) natural learning, which accounts for 10% of the total variance (Graham et al., 2002). Each of the factors significantly correlated with associated writing practices (range, $r = .17$ - $.31$). The internal consistency of the scale (i.e., alpha) is .70, demonstrating moderate reliability. Average scores for each factor were obtained, with higher scores indicating that the teacher placed more emphasis on the construct measured by that factor. Further, teachers were

also asked to answer 11 Likert-type items about their writing curriculum and 3 open-ended items about how much time their students spend writing in class. These questions were used to provide a description of the participants' writing curriculum.

Procedures

This study was conducted in three phases: (1) an eligibility and baseline assessment phase; (2) an intervention phase spanning 6 weeks; and (3) a generalization and maintenance assessment phase spanning 6 weeks. All sessions were conducted in a group format during regularly scheduled class time. Eligibility, baseline, and intervention sessions were conducted by research assistants once a week. Two generalization assessment sessions were conducted in two separate sessions during the week following the final intervention session. One generalization assessment session was conducted by research assistants, and the other was conducted by the students' primary classroom teacher. Three maintenance assessment sessions were conducted by research assistants at 2, 4, and 6 weeks following the last intervention session. Those students who were ineligible to participate in the study completed classroom instructional materials developed by the classroom teacher.

Teacher questionnaire. Prior to the eligibility and baseline assessment phase, participants' classroom teachers were asked to fill out a packet assessing their beliefs about writing instruction (i.e., Writing Orientation Scale; Graham et al., 2002), the writing curriculum they use, and their estimate of how much time their students spend writing (Appendix N). Research assistants collected the questionnaires from the teachers during the eligibility and baseline assessment phase.

Eligibility and baseline assessment phase. During the first session, students were asked to complete measures to (a) assess their eligibility to participate in the study and (b) obtain a

baseline estimate of their writing skills. To determine students' eligibility to participate in the study, the experimenter administered an informal measure of handwriting. Participants were provided with a response sheet to record their answers (see Appendix K). The examiner then read participants 10 letters from the alphabet, and instructed participants to print each letter in its lowercase form on their response sheets. Administration of this task took approximately 5 minutes. Students were considered ineligible to participate in the study if less than 90% of their letters were legible. Following this task, participants were asked to complete one training CBM-WE probe, lasting approximately 5 minutes. Results from this probe were used to provide student feedback during the subsequent session. Student participants were then administered a pre-stimulus generalization CBM-WE probe (described in detail below), lasting approximately 5 minutes. Students who wrote less than seven words on the training CBM-WE probe and the pre-stimulus generalization CBM-WE probe were considered ineligible to participate in the study. The paragraph copying task was also administered. Participants were given 90 seconds to copy a short paragraph as quickly as possible.

Participants were administered the pre-stimulus-response generalization task by their classroom teacher. Although this measure was intended to be administered prior to the first intervention session (i.e., session 2), teachers administered the measure at variable times between the baseline session (i.e., session 1) and session 3 due to scheduling conflicts.

Finally, Form A of the TOWL-3 (Hammill & Larsen, 1996) was administered in a group format after completing the CBM-WE probe during the second intervention session (i.e., week 3). Students were shown a picture, and then were given 10 minutes to plan and write a story in response to the picture prompt.

Individualized performance feedback condition. Students assigned to the experimental condition during the intervention phase received a packet containing individualized performance feedback and a CBM-WE probe. Research assistants used a procedural script to provide instructions to the students (Appendix O). The first page of the student packet contained students' identifying information. The next page of the student packet contained information regarding the individual student's performance (Appendix P). This page consisted of a box displaying the total number of words the student wrote during the previous week's session and an arrow pointing up or down. Students were told that the number in the box (i.e., total words written) was computed by counting all the words they wrote the previous week. The research assistant informed students that an upward-facing arrow indicated they wrote more words than the week prior, a downward-facing arrow indicated they wrote fewer words than the week prior, and an equal sign indicated they wrote the same amount of words as the week prior. During the first week of intervention, the number in the box displayed the total number of words written on the CBM-WE probe during the baseline phase (i.e., the CBM-WE probe from the previous week). After this step is complete, students completed a CBM-WE probe for the remainder of the session. Importantly, during the intervention phase students only received story starters that prompted them to write self-referenced narratives. Specifically, story starters only consisted of the pronoun "I" (e.g., "I opened the front door very carefully and…").

Control condition. Procedures in the control condition were exactly the same as that of the individualized performance feedback condition, but the individualized performance feedback step was omitted. Thus, student packets in this condition did not contain a performance feedback page. After listening to scripted directions from the research assistant (Appendix Q), students completed a CBM-WE probe without being informed of their progress from the previous week.

Intervention acceptability surveys. At the end of the last intervention session, student participants and their classroom teachers were asked to complete a measure of their perceptions of the intervention (see Appendices J and K, respectively). To ensure students' accurate understanding of the questions, the research assistant guided them through each question by reading them aloud.

Stimulus generalization session. To assess the extent to which students were able to generalize writing fluency gains to different stimuli, a post-intervention stimulus generalization CBM-WE probe was administered two school days after the final intervention session. This probe differed from the training CBM-WE probe in two ways: (a) rather than visually displaying a story starter at the top of the response sheet, research assistants orally presented students with a story starter and (b) rather than prompting students with a self-referenced story starter, research assistants prompted students to compose a story that contained a story stem that was not self-referenced (e.g., "As he opened the door the…"). With the exception of these probe changes and the fact that students did not receive performance feedback, all other procedures were the same as those during the intervention phase (e.g., students were given 1 minute to think about their story and 3 minutes to write).

Stimulus-response generalization session. Students were asked to complete a stimulus-response generalization task the week after the stimulus generalization task. The purpose of the stimulus-response generalization task was to examine students' ability to transfer writing fluency gains to a teacher-administered task that mimicked their typical classroom writing assignments. This task represented aspects of stimulus generalization in that it was administered by students' general classroom teachers and writing prompts were visually presented rather than visually and orally presented. Notably, writing prompts were not in story starter format. This task

incorporated aspects of response generalization in that students were expected to write a non-narrative composition (i.e., expository compare-and-contrast composition).

Maintenance sessions. Two weeks after the final intervention session, students were administered one maintenance CBM-WE probe to examine the extent to which their writing fluency gains were durable across time. With the exception of the performance feedback component, which was not included, this task was identical to intervention sessions. This maintenance session was replicated 4 weeks post-intervention and 6 weeks post-intervention.

Dependent Measures

Primary measures. Students' writing fluency progress was assessed over time by calculating the number of correct writing sequences for each CBM-WE probe. Calculating the number of correct writing sequences provided a measure of students' writing quality. Furthermore, correct writing sequences has been shown to be the most accurate measure of fluency when monitoring students' orthographic growth over time (Hubbard, 1996). Based on scoring procedures outlined by Shapiro (2004), the number of correct writing sequences was calculated by analyzing the accuracy of all adjacent words in terms of punctuation, capitalization, spelling, and syntax. Appendix R provides a detailed scoring manual.

Secondary measures. The number of total words written on each CBM-WE probe was calculated to include on individualized performance feedback forms and to indicate students' instructional level. The total number of words written was calculated by counting the total number of letter groupings separated by a space. All words were included in the total regardless of incorrect spelling. Numerals were not included in the total word count. By calculating total words written, a highly reliable measure of writing fluency, students' performance was able to be compared to national norms (Mirkin et al., 1981). Moreover, the metric of total words written

provides an indication of the grade level at which students should be instructed. Specifically, third-grade students are considered to be at the appropriate instructional level for their grade if they write 37 to 40 words in a 3-minute period (Shapiro, 2004). A lower count of total words written during that 3-minute period indicates that a student is at a frustrational level, and therefore is likely to find grade-level instruction too difficult from which to benefit. Contrarily, students who write more than 40 words during that 3-minute period are considered to have mastered third-grade level material, and therefore may benefit from higher grade-level instruction.

Further, as a description of students' initial writing abilities, standardized results of the TOWL-3 (Hammill & Larsen, 1996) are reported, and students' performance was compared to a normative national sample. Results from the teacher questionnaire are also provided to supply a further description of students' classroom writing experiences.

Experimental Design

This study used a longitudinal repeated measures design to examine students' writing fluency growth over seven weeks. Using a random number generator, all eligible student participants were randomly assigned to either the performance feedback or control condition.

Procedural Integrity

To assess the extent to which study procedures were conducted with fidelity, a permanent product measure was completed by (a) the primary research assistant for each session and (b) secondary research assistants for a subset of sessions. For each session, the primary research assistant denoted the completion of each step of the procedural script. Additionally, secondary research assistants observed 52.8% (n = 38) of the sessions. During these sessions, the secondary research assistant was equipped with a procedural script that was identical to that of the primary

research assistant. The secondary research assistant determined whether each step of the procedural script was implemented accurately, and gave credit accordingly. Agreements were scored as instances when the secondary research assistant indicated the primary research assistant accurately implemented that step. Procedural integrity was calculated by dividing the lower total count of agreements by the total number of possible procedural integrity steps and multiplying that by 100%. The mean percentage of procedural integrity recorded by the secondary research assistant was 99.9% (range, 95.8% to 100%). Table 3 details the procedural integrity calculations for both conditions.

Interscorer Agreement

A random selection of 40% of all CBM-WE probes were scored for interscorer agreement. That is, following initial data scoring, the CBM-WE probes selected for interscorer agreement were re-scored for the primary dependent measure (i.e., correct writing sequences). All instances of disagreement were re-examined by the primary researcher to make a final decision regarding the discrepancy. Interscorer agreement was calculated by dividing the number of agreements by the sum of agreements and disagreements. The mean percentage of interscorer agreement was 95.99%. To account for errors in agreement due to chance, Kappa coefficients were also calculated. The mean Kappa coefficient was .92.

Results

Data Preparation

Data input and consistency checks. The primary researcher was responsible for entering raw data into a Microsoft Excel file, which was used for its versatility in data editing. All inputted data were double-checked to attempt to increase the likelihood of accurate data entry. Data in Excel were then transferred to SPSS 16 (SPSS Inc., 2007) and SAS 9 (SAS Institute Inc.,

2002-2004). SPSS was used to compute descriptive statistics, generate graphs for data inspection, and conduct the regression analysis. A hierarchical linear modeling function in SAS was used to examine students' writing fluency progress over time. SAS was further used to conduct secondary analyses.

Data inspection. Baseline data were inspected for violations of assumptions of normality and homogeneity of variance. The assumption of normality was evaluated by calculating skewness and kurtosis. Data were considered normal if skewness was found to be within the range of +1 to -1 and kurtosis was found to be within the range of +1 to -1. Homogeneity of variance was assessed using the Levene test. Outlier data points were examined further for errors in data coding.

Descriptive analyses. Demographic data collected on the sample of participants are reported descriptively in Table 4. To determine whether differences in these demographic variables existed between students assigned to the performance feedback condition and the practice-only condition, nonparametric tests were conducted. Results indicated that no significant differences existed between conditions with regard to sex, $\chi^2(1, N = 103) = 0.01$, $p = 0.91$, race, $\chi^2(1, N = 103) = 0.06$, $p = 0.80$, ethnicity, $\chi^2 (4, N = 103) = 2.39$, $p = 0.66$, special education eligibility, $\chi^2(1, N = 103) = 0.75$, $p = 0.30$, or age, $F (1, 101) = 1.37$, $p = 0.24$. These findings suggest that on average, the demographics of students assigned to the performance feedback condition were similar to those of students assigned to the practice-only condition.

Participants' initial writing performance on three measures of writing was assessed and included: (a) Curriculum-Based Measurement in Written Expression probe, (b) the TOWL-3 Spontaneous Writing Subtest (Hammill & Larsen, 1996), and (c) the Paragraph Copying Task (Monroe & Sherman, 1966). Independent samples *t*-tests were conducted to examine whether

statistically significant differences existed between conditions on initial measures of writing performance. The students' average scores on each measure and inferential statistics results are reported in Table 5. On the Curriculum-Based Measurement in Written Expression probe, mean writing fluency of the practice-only condition was significantly higher than that of the performance feedback condition, $t(101) = 2.50$, $p = .02$. On the Spontaneous Writing Subtest of the TOWL-3, students assigned to the performance feedback condition scored significantly higher than students assigned to the practice-only condition, $t(67) = 2.14$, $p = .04$, although the scores for both groups fell within the Average range of performance. Notably, because TOWL-3 scoring guidelines require a minimum of 40 total words written for the probe to be scored, 10 students' (11%) TOWL-3 probes were not scored. No significant difference between conditions in the mean number of words copied correctly on the Paragraph Copying Task was found to exist, $t(98) = 0.05$, $p = .96$. The relationship among each of these three measures was examined, and Pearson correlation coefficients are reported in Table 6. Scores on the Paragraph Copying Task were weakly correlated with scores on the TOWL-3 Spontaneous Writing Subtest ($r = .23$, $p = .053$), but were moderately and significantly correlated with number of correct writing sequences ($r = .49$, $p < .01$) and number of total words written on the baseline Curriculum-Based Measurement probe ($r = .50$, $p < .01$). TOWL-3 scores were weakly and correlated with number of correct writing sequences ($r = .21$, $p = .08$) and number of total words written on the baseline Curriculum-Based Measurement probe ($r = .04$, $p = .78$). The number of correct writing sequences and total words written were significantly and highly correlated ($r = .94$, $p < .01$).

Classroom teachers' writing orientations and classroom instructional practices. Teachers were asked to complete a questionnaire designed to examine their orientations to writing instruction (Graham et al., 2002). Results of this measure indicated that the teachers

placed the most emphasis on explicit instruction (*M* factor score = 4.67; *SD* = 1.13) and natural writing (*M* factor score = 4.29; *SD* = 0.95), while they placed slightly less emphasis on correct writing (*M* factor score = 3.03; *SD* = 1.25; see Table 7). Four different writing curricula or techniques (i.e., McGraw-Hill New Treasures Program, Lucy Calkins Program, Four-Square Method, and graphic organizers) were reportedly used by the teachers to inform their classroom instruction. The teachers reported spending most of their instructional time each week on spelling practice (*M* = 55.8 minutes, *SD* = 30.4 minutes) and composition practice (*M* = 52.5 minutes, *SD* = 34.9 minutes).Less time was allocated for weekly handwriting practice (*M* = 30 minutes, *SD* = 36.9 minutes). When asked to estimate the frequency of their own specific instructional writing practices, the following were reported to be used daily by some teachers: (a) invented spelling; (b) specifically teaching grammar skills; (c) re-teaching writing skills and strategies; (d) specifically teaching spelling skills, and (e) specifically teaching handwriting skills (see Table 8).

Major Analyses

Analyses were conducted to assess (a) whether performance feedback significantly improved students' writing fluency over time, (b) whether students were able to maintain gains in writing fluency over a period of time, and (c) whether the intervention resulted in students' ability to generalize their writing skills to different writing formats.

Performance feedback. The trajectory of students' writing fluency growth throughout the duration of the intervention phase was examined for the performance feedback and control conditions. Students' growth in writing fluency (i.e., slope) was computed using the metric of correct writing sequences. Multilevel modeling was used to analyze whether a statistically significant difference in the slope of writing growth existed between conditions. These between-

condition differences in the trajectory of student participants' writing growth over time were examined by using a mixed-model repeated measures design (PROC MIXED function in SAS V9.2 software, SAS Institute, 2000). Level 1 and Level 2 analyses were specified to evaluate the first primary research hypothesis. The Level 1 analysis was used to estimate the patterns of intra-individual growth by a linear model, which contained both intercept (i.e., estimated baseline performance) and slope (i.e., rate of change in performance across sessions). The Level 2 analysis then was used to examine the between-condition differences in the intercept and slope.

First, an empty model that contained only the intercept was analyzed. The intraclass correlation (ICC), which is a measure of within-person variability, was calculated by hand using the intercept and residual estimates that were produced by the empty model. Results indicated that approximately 50.29% of the total variance in this model was explained by within-person variability. According to guidelines put forth by Lee (2000), these results support the use of multilevel modeling as an appropriate analysis for these data.

Intervention session was added to the empty model to produce the Level 1 model (i.e., unconditional model). The addition of the time variable (i.e., session) to the model accounted for a substantial amount of variance (*pseudo* $R^2 = 0.22$). Results from this model suggest that participants demonstrated significant gains in correct writing sequences across intervention sessions, with an average gain of 1.35 correct writing sequences per session, $t\ (102) = 6.30$, $p < .001$.

To determine if additional variables should be included in the final conditional growth model, the predictor variables of handedness, sex, and baseline instructional level (i.e., frustrational, instructional, mastery) were first entered into the model for each individual. The percentage of variance explained in the conditional growth model by each of these variables was

examined. Adding the predictor of sex made an insignificant contribution to the model, t (439) = -0.58, p = .56, as only a very small amount of variance in intercept (*pseudo* R^2 = 0.06) and slope (*pseudo* R^2 = 0.009) was explained by this predictor. A similar result was obtained upon adding the handedness variable as a predictor, t (399) = -0.90, p = .37, which resulted in the explanation of only 5% variance in intercept (*pseudo* R^2 = -0.05) and 4.8% variance in slope (*pseudo* R^2 = 0.048). Conversely, adding students' baseline instructional level to the model explained a statistically significant amount of variance in the intercept (*pseudo* R^2 = .558) and slope (*pseudo* R^2 = .169), t (439) = -2.82, p = .005. Thus, students' baseline instructional level was included as a covariate in the final conditional growth model (i.e., Level 2).

Results of the final conditional growth model revealed that, after controlling for baseline instructional level, students assigned to the performance feedback condition evidenced statistically significantly greater writing fluency growth over time than those assigned to the practice-only condition, t (439) = 10.72, p < .001,d = 0.89, CI_{95}[0.64, 1.13]. Parameter estimates were examined as a function of baseline instructional level based on the final conditional growth model. Because too few participants wrote at the instructional level at baseline (n = 3), parameter estimates could not be calculated for this group. Thus, parameter estimates are reported only for those who wrote at the frustrational and mastery levels at baseline.

For students who performed within the frustrational level (i.e., wrote fewer than 37 words per 3 minutes) at baseline, a statistically significant difference in slopes existed between the conditions in favor of those who received the performance feedback intervention, t (401) = 11.06, p < .001, d = 0.94, CI_{95} [0.89, 1.31]. Students who were writing at the frustrational level at baseline and were assigned to the practice-only condition gained an average of 0.35 correct writing sequences per week. In contrast, students who were writing at the frustrational level at

baseline and were assigned to the performance feedback condition gained an average of 2.62 correct writing sequences per week (see Figure 4).

A statistically significant difference in slopes between conditions was also found for students writing at the mastery level (i.e., wrote more than 40 words per 3 minutes) at baseline, $t(25) = 4.22$, $p = .05$, $d = 1.75$, CI_{95} [0.73, 2.54]. For those students assigned to the performance feedback condition, an average of 0.93 correct writing sequences was gained per week. However, the students assigned to the practice-only condition demonstrated an average decline of 2.21 correct writing sequences per week (see Figure 5).

Generalization. To examine whether students in the performance feedback condition differed significantly from students in the practice-only condition on measures of generalization, one-tailed analyses of covariance (ANCOVA) were computed with alpha set to .05.

The first ANCOVA was conducted to determine whether students assigned to the performance feedback condition evidenced significantly greater levels of writing fluency on the stimulus generalization measure than students assigned to the practice-only condition. Prior to conducting the ANCOVA, its underlying theoretical assumptions were examined. First, the covariate (i.e., CWS on the pre-stimulus generalization measure) was found to be significantly and moderately correlated with the dependent variable (i.e., CWS on the post-stimulus-response generalization measure), $r = .55$, $p < .001$. As recommended by Tabachnick and Fidell (2007), a scatterplot was created to evaluate the assumption of linearity between the covariate and the dependent variable. Visual inspection of the scatterplot indicated that the relationship between the covariate and the dependent variable was linear. Next, the assumption of homogeneity of regression slopes was analyzed to ensure that no interaction existed between the covariate and the condition to which the participants were assigned. Results from a univariate analysis of

variance revealed that this assumption was also met, F (1, 87) = .01, p = .94. Finally, Levene's test of equality of error variances indicated that no significant difference in error variances existed between groups (i.e., homogeneity of variance), F = 2.21, p = .14. Because all assumptions were met, it was determined that an ANCOVA was an appropriate statistic to use in this case.

Table 9 lists the results of the ANCOVA, which suggest that when controlling for baseline fluency on the pre-stimulus generalization probe, students assigned to the performance feedback condition had significantly more correct writing sequences on the post-stimulus generalization probe than those assigned to the practice-only condition. Furthermore, while writing fluency decreased from pre- to post-assessment for those assigned to the practice-only condition, it increased for the students who received performance feedback. These results indicate that assignment to the performance feedback condition resulted in significantly greater generalization to a task that differed from the intervention.

To examine the extent to which student participants were able to generalize intervention gains to a writing task that mimicked a typical classroom writing assignment (i.e., stimulus-response generalization), another one-way ANCOVA was conducted. Assumptions of the ANCOVA were again examined to ensure the appropriateness of this analysis. The covariate (i.e., correct writing sequences on the pre-stimulus-response generalization measure) was significantly and moderately correlated with the dependent variable (i.e., correct writing sequences on the post-stimulus-response generalization measure), r = .56, p < .001. Visual inspection of the scatterplot indicated that the assumption of linearity between the covariate and dependent variable was also upheld. Similarly, the assumption of homogeneity of regression slopes was also met, F (1, 87) = .001, p = .99. A Levene's test indicated that the error variances

between the groups were not significantly different (i.e., homogeneity of variance), although it approached statistical significance, $F = 3.82$, $p = .054$. Therefore, the second ANCOVA was also deemed appropriate given these data.

Results of the second ANCOVA are presented in Table 10. No significant difference was found to exist in the number of correct writing sequences as a function of group assignment when controlling for baseline performance, and the effect of condition on writing fluency generalization was small, partial $\eta^2 = .004$. Interestingly, the trend in the data for the stimulus-response-generalization task was opposite that of the stimulus generalization task. That is, students assigned to the performance feedback condition actually demonstrated a decline in the number of correct writing sequences on this task, whereas those students assigned to the practice-only condition demonstrated improvement. These results suggest that assignment to the performance feedback condition did not result in the ability to generalize writing gains on a task that resembled students' typical classroom writing assignments.

Maintenance. To assess the maintenance of intervention effects, the percentage of maintenance gains and/or losses was calculated. Specifically, for each individual, the percent gain or loss was calculated by subtracting the score on the final training probe from the score on the maintenance probes, dividing this number by the score on the final training probe, and multiplying by 100. The mean percent change was then obtained for each condition on (a) the 2-week maintenance CBM probe, (b) the 4-week maintenance CBM probe, and (c) the 6-week maintenance CBM probe. To determine whether the percent change on each maintenance probe was significantly different as a function of condition, three independent samples *t*-tests were conducted. For each *t*-test, the assumption of homogeneity of variance was evaluated by computing a Levene's test. This assumption was supported for the 2-week maintenance data, $F =$

.78, $p = .38$; however, Levene's test indicated unequal variances for the 4-week, $F = 10.97$, $p = .001$, and 6-week, $F = 7.82$, $p = .007$, maintenance sessions. Therefore, alternative *t* values were computed to compensate for the unequal variances, and degrees of freedom were adjusted accordingly. Descriptive statistics, percent change, *t*-test, and effect size results are reported in Table 11.

The percent change results revealed that students assigned to the performance feedback condition evidenced a slight gain (1.58%) in writing fluency from the final intervention session to the first maintenance session. However, students assigned to the practice-only condition evidenced a 31.8% gain on the 2-week maintenance probe, suggesting that they were able to maintain intervention effects to a significantly greater extent than students assigned to the performance feedback condition.

On the 4-week maintenance probe, the performance feedback condition demonstrated a 10.56% loss in writing fluency, while the practice-only condition continued to improve (42.93% gain). Between-conditions analysis again indicated that students assigned to the practice-only condition maintained fluency gains to a significantly greater extent than students assigned to the performance feedback condition at 4 weeks post-intervention.

To further aid in the analysis of maintenance, Pearson's correlation coefficients were computed to determine the extent to which students' performance on maintenance probes related to their performance on training probes from the final intervention session. For each condition, the correlations between the mean correct writing sequences on the final training CBM probe and each maintenance probe were calculated. Correlation results appear in Table 12. For the performance feedback condition, mean correct writing sequences at the final intervention session were highly and significantly correlated with their mean correct writing sequences at each

maintenance session (r = .82, .69, .76). Although statistically significant, the practice-only condition's mean correct writing sequences at the final intervention session were not as strongly correlated with their performance at any of the maintenance sessions (r = .57, .47, .35). Thus, these findings suggest that in comparison to the practice-only control condition, the performance feedback condition's writing fluency performance at each maintenance session was more highly related to their performance at the final intervention session.

Secondary Analyses

Secondary analyses were conducted to examine (a) whether there was a statistically significant shift in students' instructional level classification from baseline to post-intervention and (b) factors that predict students' maintenance and generalization.

Instructional level. A McNemar-Bowker test was used to examine whether the percentage of students at each instructional level (i.e., frustrational, instructional, and mastery) changed significantly from pre-intervention to post-intervention. Results indicated that across both conditions, shifts in instructional level was significant following intervention, $\chi_{MB}^2 = 27.80$, df = 3, $p < .001$. When each condition was analyzed independently, results indicated that although the performance feedback condition experienced significant shifts in instructional level from pre- to post-intervention, $\chi_{MB}^2 = 29.0$, df = 3, $p < .001$, the practice-only condition did not, $\chi_{MB}^2 = 1.95$, df = 2, $p = .38$.

Table 13 displays the percentage of students classified at each instructional level for the total sample and by condition. At baseline, the majority (92.2%) of all students performed within the frustrational range of writing fluency. Upon completion of the intervention phase, over half (51.1%) of the students assigned to the performance feedback condition reached the mastery level (i.e., wrote at least 40 words per 3 minutes). In contrast, only 13% of the students

assigned to the practice-only condition performed within the mastery range at the final session. Positive shifts toward the instructional level also occurred for the performance feedback condition, with 14.9% of students assigned to that condition writing within the instructional range during the final intervention session. A smaller percentage (8.7%) of students in the practice-only condition performed within the instructional range at the final session, but this percentage was still higher than baseline levels. Upon conclusion of the intervention phase, 34% of students assigned to performance feedback condition and 78.3% of students assigned to the practice-only condition continued to write at the frustrational level.

Predictors of maintenance and generalization. To examine factors that significantly predicted maintenance and generalization, three multiple regression analyses were proposed to be conducted. One multiple regression aimed to examine the predictive impact of intervention factors (i.e., instructional level at baseline, instructional level at final intervention session, and slope) and student demographic factors (i.e., absenteeism, ethnicity, and acceptability of intervention score) on maintenance (i.e., difference between correct writing sequences of final training CBM probe and the 6-week maintenance probe). Two additional multiple regression analyses aimed to assess the impact of the same predictor variables on students' ability to generalize writing fluency skills to (a) the stimulus generalization probe and (b) the stimulus-response generalization probe.

Prior to conducting the multiple regression analyses, the statistical assumptions were examined according to guidelines described by Tabachnick and Fidell (2007). Because the skewness and kurtosis were close to 0 for each of the three dependent variables, sample size requirements were evaluated, and the results indicated that the obtained sample (N = 73 for maintenance, 87 for stimulus generalization, and 88 for stimulus-response generalization) was

not of adequate size. In addition, upon evaluating each of the independent variables for normality, it was found that several independent variables had high skewness (i.e., absenteeism, baseline instructional level; skew range = 1.7 to 3.6) and kurtosis (i.e., baseline instructional level, slope, absenteeism; kurtosis range = 2.12 to 11.79) values. Analysis of multicollinearity indicated that although none of the independent variables were highly correlated with each other, only two of the independent variables (i.e., mastery level at the final intervention session and slope) were even moderately correlated with the maintenance and stimulus generalization variables. Furthermore, all independent variables were weakly correlated with the stimulus-response generalization variable. Visual inspection of scatterplots and the normal probability plots of the regression standardized residuals indicated that (a) the distribution of the dependent variables did not significantly deviate from normality, (b) the relationship between the predicted dependent variable scores and the residuals was linear, and (c) the variability of the residuals was homoscedastic. As a result of this examination, multiple regression analyses were not conducted because a significant number of the statistical assumptions were violated.

Student acceptability of intervention procedures. The overall acceptability of intervention procedures was moderately acceptable for students assigned to both the performance feedback condition (M = 4.14, SD = 1.25) and the practice-only condition (M = 3.91, SD = 1.39; see Table 14). An independent samples t-test was conducted to examine whether significant differences existed in students' perceived overall acceptability as a function of the condition to which they were assigned. Results showed that there was not a significant difference between conditions in students' perceptions of procedures, t (90) = 0.83, p = .41.

Teacher acceptability of intervention procedures. Overall, the student participants' general education teachers perceived the intervention procedures to be moderately acceptable (M

= 4.7, *SD* = 0.92). See Table 15 for a summary of the teachers' perceptions of intervention procedures. Aspects of the intervention that teachers described as highly acceptable include (a) that students' writing difficulties are severe enough to warrant its use (*M* = 5.50, *SD* = 0.55), (b) that the intervention would not likely result in negative side effects (*M* = 5.17, *SD* = 0.41), and (c) the intervention's procedures (*M* = 5.17, *SD* = 0.75). However, teachers' reports indicate that they had less confidence in the intervention's effectiveness (*M* = 4.33, *SD* = 0.52) and suitability for the writing difficulties of their students (*M* = 4.33, *SD* = 0.52). Furthermore, they indicated that this type of intervention was fairly inconsistent with procedures they have used previously (*M* = 3.50, *SD* = 2.07).

Discussion

Overall, results of this study support previous findings that suggest providing students with weekly performance feedback significantly increases growth in writing fluency over the course of a 6-week intervention. Furthermore, writing growth for students assigned to the performance feedback condition significantly exceeded the growth that was displayed by students assigned to the practice-only condition. In addition to greater evidence of writing fluency growth over time, students assigned to the performance feedback condition experienced positive shifts in instructional level that reached statistical significance. In fact, at the final intervention session, 51.1% of students who received performance feedback had reached the mastery level of writing fluency. A similar pattern of shifts was not observed within the practice-only condition; for these students, a substantial percentage (78.3%) continued writing in the frustrational range at the conclusion of the intervention.

Despite the finding that students assigned to the performance feedback condition were better able to acquire the ability to write with fluency, evidence of maintenance and

generalization of this skill was limited. Specifically, students assigned to the performance feedback condition demonstrated evidence of stimulus generalization to a significantly greater extent than those assigned to the practice-only condition, but they failed to generalize those skills to a task that was similar to their typical classroom writing assignments. Additionally, students assigned to the practice-only condition demonstrated significantly greater evidence of maintenance at each follow-up session than students assigned to the performance feedback condition, which was directly contrary to the initial hypothesis. The primary findings of this study are discussed in further detail below.

Effectiveness of Performance Feedback versus Practice-Only in Improving Students' Writing Fluency over Time

Practice and performance feedback have both been shown to improve students' writing fluency (Eckert et al., 2006; Harris et al., 1994). Due to this finding, students in both conditions (i.e., performance feedback and practice-only) were expected to make gains in writing fluency growth over time. However, the first primary hypothesis of this study was that those students assigned to the performance feedback condition would demonstrate fluency growth to a significantly greater extent than those assigned to the practice-only condition. Indeed, this hypothesis was confirmed by the present study's results; although all students in the study demonstrated an average gain of 1.35 correct writing sequences per session, the growth trajectory of correct writing sequences for students assigned to the performance feedback condition was significantly steeper than that of the practice-only condition. The performance feedback intervention appeared to be particularly beneficial for students who were writing at the frustrational level at baseline. These students gained an average of 2.62 correct writing sequences per week, whereas students writing at the frustrational level at baseline who received

practice alone for 6 weeks gained an average of only 0.35 correct writing sequences per week. This finding is comparable to national reporting standards suggesting that third-grade students who receive typical classroom instruction alone (i.e., do not receive formal writing intervention) gain an average of 0.30 correct writing sequences per week (AIMSweb, 2010).

Students who were writing at the mastery level at baseline did not appear to derive as much benefit from the performance feedback intervention. These students gained an average of 0.93 correct writing sequences per week, which is still higher than what would be expected given typical classroom instruction alone. Conversely, the writing fluency of students who were writing at the mastery level at baseline and received practice alone for 6 weeks actually decreased by an average of 2.21 correct writing sequences per week. In future studies, it may be useful to conduct follow-up interviews with these students in an attempt to determine which factors affect their writing over time.

Effectiveness of Performance Feedback versus Practice-Only in Producing Generalized Effects of Writing Improvement

The task force that guides evidence-based intervention knowledge within the field of school psychology (Kratochwill & Stoiber, 2002) recommended that all treatment studies should examine the extent to which participants are able to generalize intervention effects across settings and persons. To adhere to this guideline, evidence of stimulus generalization and stimulus-response generalization was examined in this study. For both generalization processes, it was hypothesized that students assigned to the performance feedback condition would outperform those assigned to the practice-only condition.

In an attempt to assess stimulus generalization, an alternative measure was created that differed from the training measure only in its presentation of the story-starter. Specifically, the

story-starters for the training measure were both visually and orally presented and contained only self-referenced pronouns (i.e., I), whereas the story-starters for the stimulus generalization measure were only orally presented and contained only other-referenced pronouns (i.e., she, he). Results supported the hypothesis that students assigned to the performance feedback condition did in fact transfer writing gains to the stimulus generalization task to a significantly greater extent than those students assigned to the practice-only condition. In fact, although students assigned to the performance feedback condition gained an average of 3 correct writing sequences on the stimulus generalization measure from pre-intervention to post-intervention, students assigned to the practice-only condition lost an average of 3 correct writing sequences.

In an attempt to assess stimulus-response generalization, an alternative measure was created that consisted of a task that mimicked not only students' typical classroom writing assignments, but also their end-of-year writing test. Thus, generalization to this measure represented a clinically meaningful transfer of intervention effects. The hypothesis that students assigned to the performance feedback condition would demonstrate significantly greater evidence of stimulus-response generalization than those assigned to the practice-only condition was not supported by the results of this study. Rather, results revealed that there was not a statistically significant difference between conditions on the post-stimulus-response generalization measure when controlling for baseline writing fluency. In fact, the trend in the data showed that writing fluency decreased from pre- to post-intervention on this measure for students who received 6 weeks of performance feedback, whereas writing fluency increased for students who received practice only.

There are several possible reasons that the stimulus-response generalization hypothesis was not supported by results of this study. First, methodological problems could have

confounded results. Although the pre-stimulus-response generalization measure was supposed to be administered by the students' general education teachers during the baseline phase (i.e., prior to the first intervention session), many of the teachers were unable to do so due to scheduling conflicts. As a result, students were administered this task at variable times between the second and third intervention sessions as a function of their classroom teachers. This irregularity could have impacted results, as later administration of the pre-stimulus-response generalization measure would likely lead to higher "baseline" scores for the performance feedback condition. Furthermore, although teachers were provided with a procedural script for this task, procedural integrity data were not collected because definite times of administration were not identified by the teachers. Due to this, the extent to which all procedures were followed by the teachers is unknown. Based on anecdotal information gathered, it was determined that at least one major administration error occurred during the pre-stimulus-response generalization session. Specifically, one teacher reported that she orally presented the essay topic, which was supposed to be visually presented only. Finally, teachers' behaviors prior to this task were not controlled, so some teachers may have provided task-specific instruction beforehand that could have influenced results and limited the internal validity of this assessment.

A second possible reason that there was no statistically significant difference between conditions on the stimulus-response generalization measure was that the stimuli and responses required for this task may have differed too much from the training sessions. For example, specific stimuli that differed significantly between the training and stimulus-response generalization sessions included the task administrator, the presentation of the writing topic, and the length of the planning and writing periods. The writing response required for this task also

significantly differed from the training sessions. Although the training session probes required students to write narratives, this task required students to write compare-and-contrast essays.

A final factor that may account for the lack of stimulus-response generalization is the length of the writing period for this measure, as research in the occupational therapy field has shown that writing for as little as 10 minutes can produce handwriting fatigue (e.g., Parush, Pindak, Hahn-Markowitz, & Mazor-Karsenty, 1998). Although students wrote for only 3 minutes on training probes, they were expected to write for 10 minutes on the stimulus-response generalization task. Previous research with college students and children has found that handwriting speed on shorter tasks does not predict handwriting speed on longer tasks (O'Mahony, Dempsey, & Killeen, 2008; Summers & Catarro, 2003) and that handwriting speed significantly decreased over a 10 minute writing period (Kushki, Schwellnus, Ilyas, & Chau, 2011). It has been suggested that biomechanic changes that are associated with handwriting fatigue (e.g., limb stiffness) may be a contributing factor.

Effectiveness of Performance Feedback versus Practice-Only in Maintaining Students' Writing Fluency Gains over Time

A final primary hypothesis of this study was that students assigned to the performance feedback condition would maintain writing gains over a 2-, 4-, and 6-week period to a significantly greater extent than students assigned to the practice-only condition. This hypothesis was not supported by the results of this study, which indicated that at each maintenance session, students assigned to the practice-only condition evidenced a significantly higher percent change from the final intervention session than students assigned to the performance feedback condition. However, writing fluency on each maintenance probe was more highly related to writing fluency on the final intervention probe for students assigned to the

performance feedback condition. This is further evidenced by the high percent gains of correct writing sequences for the practice-only condition (i.e., 31.8%, 42.9%, and 21.44%) and the lower percent gains/losses for the performance feedback condition (i.e., 1.6%, -10.6%, and -5.3%). Additionally, the mean number of correct writing sequences for the performance feedback condition (M = 31.62) was much higher than that of the practice-only condition (M = 23.59) at the final intervention session. As a result, despite having a 5.3% loss, the performance feedback condition still had higher mean correct writing sequences at the 6-week maintenance session than the practice-only condition (M = 28.26, M = 24.33, respectively), which had a 21.4% gain. Thus, although the percent change scores for each maintenance session reflect continued writing fluency gains for the students assigned to the practice-only condition, students assigned to the performance feedback condition demonstrated fluency performance that was more consistent with their performance on the final intervention probe.

One possible explanation for why students assigned to the practice-only condition continued to demonstrate growth in writing fluency over the maintenance phase concerns the administration procedure that was used. For students assigned to the practice-only condition, the maintenance phase procedures were the same as those that they had experienced during the intervention phase. However, a salient stimulus was missing for students assigned to performance feedback condition. Those students received feedback during the intervention phase, but did not during the maintenance phase. Therefore, while the presentation of task materials was different for students assigned to the performance feedback condition, it remained constant for students assigned to the practice-only condition. Given those conditions, it is not surprising that students assigned to the practice-only condition continued making fluency gains

while students assigned to the performance feedback condition demonstrated some evidence of decline.

A second possible explanation for the results observed in the maintenance data pertains to a literature base in the field of organizational behavior management. Specifically, Agnew and Redmon (1992) theorized that feedback may act to prompt individuals to create a rule, or a contingency-specifying stimulus (i.e., an "if…then" statement that specifies a behavior and a consequence), regarding their own performance. This, in turn, may alter the function of other stimuli such that they become discriminative stimuli or reinforcing stimuli and thus serve to change behavior. In this sense, it is possible that students who received feedback developed a rule for themselves, such as, "If I write more words, then I will receive an upwards-facing arrow, and my peers may congratulate me on my performance." This type of contingency-specifying self-talk could lead to an alteration in the function of stimuli associated with writing fluency (e.g., the writing packet could become a discriminative stimulus that affects the number of words written, and performance feedback could become a reinforcing stimulus). According to this theoretical framework, the provision of performance feedback may have maintained high rates of student responding during the training sessions, but without the presence of this function-altered reinforcing stimulus during the maintenance sessions, those high rates of text production were not sustained.

Limitations

Several methodological aspects of this study may limit one's confidence in the results. One such important limitation of this study involved the extent to which procedures were followed as intended for the stimulus-response generalization session. Because researchers were not present for this teacher-administered task, procedural adherence during these sessions is

unknown. Furthermore, one teacher indicated that she orally provided students with the writing prompt during the pre-stimulus-response generalization session despite the fact that the procedural instructions required her not to do so.

Another problem associated with the stimulus-response generalization measure was the timing of administration. Although teachers were supposed to administer this measure during the baseline phase, due to scheduling difficulties it was not administered until the second and third sessions of the intervention phase. Furthermore, due to variability in teacher's schedules, two teachers administered this measure after the students received two sessions of performance feedback, and the remaining four teachers administered it after the students received three sessions of performance feedback. Also concerning procedural integrity, a subset of students in the performance feedback condition did not receive an oral reminder of the story-starter prompt during the post-stimulus generalization session. This could have reduced the mean writing fluency score for the performance feedback condition, especially considering that students were not visually provided with the story-starter during that session.

Another threat to the internal validity of this study is the threat of selection error. Although all students were randomly assigned to one of the two conditions, and thus potential selection biases should have been eliminated, the mean number of correct writing sequences on the baseline training probe was significantly higher for the practice-only condition than the performance feedback condition.

Although this study sought to examine the effects of a performance feedback intervention with a typically developing population, question remains as to how "typical" this sample of students was. Demographic data revealed that a large proportion of students received free or reduced-priced lunch and represented minority ethnic backgrounds. Further, because the student

sample consisted only of third-grade students from an urban school setting, the extent to which the results of this study can be extrapolated to the general, elementary-aged student population in the United States is limited. Thus, threats to external validity were also present in this study.

Directions for Future Research

Previous research (Cutler & Graham, 2008; National Commission on Writing, 2003) has suggested that the significant national deficit in children's writing performance reflects the broader problem of a lack of effective teaching procedures. Given this, these authors urged for the further development and use of evidence-based teaching practices for writing in the general education setting. Because the performance feedback intervention described in this study (a) significantly increased students' writing fluency growth and (b) can be administered in a class-wide format, it is not inconceivable that this intervention may be considered by some to be an effective method for writing fluency instruction within the general education classroom. However, although this performance feedback intervention tended to be effective for many students, some results of this study suggest that in its current state, this class-wide intervention may not be an ideal instructional technique for third-grade general education classrooms. First, 66% of students who received performance feedback for 6 weeks finished the intervention by writing at the instructional or mastery level. Ideally, effective instruction should result in a higher percentage of students writing at these levels. Second, effective instructional methods should maintain writing fluency growth over time and result in the ability to transfer gains to meaningful writing tasks. Currently, this class-wide intervention does not fulfill these requirements. Indeed, students assigned to the performance feedback condition were able to transfer fluency gains to a task that differed slightly in presentation of the story-starter (i.e., oral rather than visual) and in the pronouns used in the story-starter (i.e., other-referenced rather than

self-referenced). Nonetheless, they were unable to transfer these gains to the stimulus-response generalization measure, a task that was inarguably more clinically meaningful than the stimulus generalization measure.

Arguably the most pressing next stage in this research given the results of the current study is to take steps to systematically program for generality during the training sessions. This could be achieved by analyzing the programming techniques described by Stokes and Osnes (1989) and ensuring that the most salient aspects of the study have been programmed for generality. One such aspect that future researchers should seek to refine is the training material. In the current study, CBM-WE probes were used for training, stimulus generalization, and maintenance sessions. Students in the performance feedback condition demonstrated evidence of stimulus generalization and maintenance at 2 weeks post-intervention. However, their performance on a measure that was structurally dissimilar from the training probes (i.e., stimulus-response generalization measure) decreased from pre- to post-intervention. As a general principle, stimuli that are more similar to training materials will be more likely to evoke a generalized response than stimuli that are different (Cooper et al., 1987). Thus, if generalization to the stimulus-response generalization probe is of the most clinical importance due to its representation of real-world conditions, it may be beneficial to ensure that the training probes more closely resemble the stimulus-response generalization probe.

In addition to programming for generality in a more systematic manner, there is also a need to refine the performance feedback intervention so that it produces more salient gains among students who are already writing at the mastery level at baseline. Although these students evidenced writing growth as a result of performance feedback in the current study, they gained fewer correct writing sequences per week than students who were writing at the frustrational

level at baseline. Of course, this may be a function of the possibility that a fluency-based intervention may be inappropriate for students who are already fluent writers. Thus, future research should aim to identify aspects that could be incorporated into the performance feedback intervention to make it more salient for those children who begin at the mastery level at baseline. Research that has examined performance feedback among "experts" has been conducted with a college-aged sample. Results showed that students who had high prior knowledge about a task tended to benefit not from individual feedback alone, but from the opportunity to work collaboratively with other "experts" and discuss the learning task and possible solutions (Nihalani, Mayrath, & Robinson, 2011). Therefore, incorporating these types of opportunities in combination with aspects of the performance feedback intervention may be a fruitful area for researchers.

A final consideration for future research is to explore classroom teachers' writing orientations and teaching procedures that predict students' responsiveness to the performance feedback intervention and the likelihood that intervention effects will generalize and maintain. Theoretical conceptualizations of writing have suggested that handwriting is an important skill that contributes to writing fluency in elementary-aged children (Berninger et al., 1997; Graham et al., 1997). Therefore, the amount of time that teachers spend on handwriting instruction may be a significant predictor of students' ability to acquire, maintain, and generalize fluency gains. Additionally, research should examine whether students who receive little direct instruction in handwriting are more likely to begin the intervention by writing at the frustrational level at baseline. In the current study, teachers reported allotting a small percentage of time (M = 30 minutes per week) to handwriting instruction, and most teachers reported that they did

specifically teach handwriting skills more than monthly. This may aid in the explanation of why so many students began the intervention writing at the frustrational level.

Conclusion

Currently, given the substantial percentage of elementary-aged students who perform below grade-level on measures of writing ability, there is a need for empirically-validated interventions that directly address this concern (National Commission on Writing, 2003). One such intervention involves the provision of feedback regarding students' performance on measures of writing fluency. Although performance feedback has been shown to significantly improve elementary-aged students' writing fluency in comparison to receiving practice alone (Eckert et al., 2006), no studies to date have assessed the extent to which students demonstrate evidence of generalization and maintenance of these writing gains. Thus, the current study sought to replicate previous research that examined performance feedback in the context of developing writing fluency skills, and to extend that research by examining students' ability to maintain and generalize intervention effects.

Results of this study revealed that students who received weekly, individualized performance feedback demonstrated significantly greater growth in writing fluency over time than students who received weekly writing practice alone. In comparison to students assigned to the practice-only condition, these students also demonstrated significantly greater evidence of generalization to a task that differed from the training sessions in terms of the format of the stimulus material (i.e., stimulus generalization). No significant difference was found to exist between conditions on a generalization measure that mimicked students' typical classwork (i.e., stimulus-response generalization), although limitations associated with the administration of this measure could explain the students' failure to show evidence of this type of generalization.

Finally, while students assigned to the practice-only condition demonstrated significantly greater evidence of maintenance of intervention effects across 6 weeks post-intervention, the number of correct writing sequences on each maintenance probe was more highly related to performance on the final intervention session for students assigned to the performance feedback condition. Given the results of this study, future research should examine the effects of systemically programming generality, incorporating aspects that may be salient to students who write at the mastery level at baseline, and instructional predictors of intervention effects and generality.

Table 1

Story Starter Prompts

1. I opened the front door very carefully and…

2. I once had a magic pencil and…

3. I was chewing a piece of bubble gum when…

4. I was on my way home from school and...

5. I was talking to my friends when all of a sudden...

6. One day I found the most interesting thing...

7. One night I had a strange dream about...

8. I was playing outside when a spaceship landed and…

9. One day I went for an airplane ride and…

10. I was in the middle of the lake when…

11. She woke from a sound sleep when something…

12. As he opened the door the…

Table 2

Studies Examining the Validity and Reliability of Curriculum-Based Measurement in Written Expression

Study	Grade Level	Metric	Criterion Measure	Validity Coefficient	Reliability Type	Reliability Measure
Deno, Mirkin, & Marston (1980)	3 to 6	TWW CSW	TOWL	.41 - .82 .45 - .88		
Marston & Deno (1981) – Study 1	1 to 6	TWW CSW			Parallel Forms	.95 .95
Marston & Deno (1981) - Study 2	1 to 6	TWW CSW			Split Half	.99 .96
Videen, Deno, & Marston (1982)	3to 6	CWS	DSS TOWL Holistic rating	.49 .69 .85	Interscorer	.90
Tindal, Germann, & Deno (1983)	4	TWW			Parallel Form	.70
Shinn, Ysseldyke, Deno, & Tindal (1982)	1 to 5	TWW			Parallel Form	.51 - .71
Fuchs, Deno, & Marston (1982)	3 to 6	CSW			Parallel Form	.55 - .89
Marston, Deno, & Tindal (1983)	3 to 6	TWW CSW			Interscorer	.96 .91
Tindal, Martson, & Deno (1983)	1 to 6	TWW CSW			Parallel Form	.73 .72

Study	Grade Level	Metric	Criterion Measure	Validity Coefficient	Reliability Type	Reliability Measure	
Tindal & Parker (1991)	3 to 5	TWW CSW CWS	Stanford	.18 - .25 .22 - .30 .31 - .41			
Parker, Tindal, & Hasbrouk (1991)	2 to 5	TWW CSW CWS	Holistic rating	.36 - .49 .43 - .64 .58 - .61			
Gansle, Noell, VanDerHeyden, Naquin, & Slider (2002)	3 to 4	TWW CSW CWS	Teacher Ratings	.08 .21 .36	Parallel Form & Interscorer	.62 .53 .46	.96 .95 .86
Gansle, Noell, VanDerHeyden, Slider, Hoffpauir et al. (2004)	3 to 4	TWW CWS	WJ-R Writing Samples	.23 .36			
Malecki & Jewell (2003)	1 to 8	TWW CSW CWS			Interscorer	.99 .99 .98	

Note. TWW = Total Words Written, CSW = Correctly Spelled Words, CWS = Correct Writing Sequence

Table 3

Descriptive Statistics for Procedural Integrity Assessments

	Sessions Assessed		Integrity Results		
Phase/Condition	%	(*n*)	*X*	*SD*	Range
Baseline	50	3	100	0	--
Practice-Only	72.2	13	100	0	--
Feedback	66.7	12	100	0	--
Post-Stimulus Generalization	33.3	2	97.9	2.9	95.8-100
Stimulus-Response Generalization (Pre and Post)	0	0	--	--	--
2-Week Maintenance	33.3	2	100	0	--
4-Week Maintenance	66.7	4	100	0	--
6-Week Maintenance	33.3	2	100	0	--
Overall	52.8	38	99.9	0.7	95.8-100

Notes. Baseline procedural integrity assessment contained 45 steps. Practice-only procedural integrity assessment contained between 21 and 45 steps; feedback procedural integrity assessment contained between 24 and 53 steps; post-stimulus generalization procedural integrity assessment contained 24 steps; maintenance procedural integrity assessment contained 19 steps.

Table 4

Student Demographic Information (N = 103)

	Condition							
	Total Sample		Practice-Only		Performance Feedback			
Characteristics	%	(*n*)	%	(*n*)	%	(*n*)	χ^2	*p*
Sex							0.01	.91
Female	49.5	(51)	23.3	(24)	22.3	(23)		
Male	50.5	(52)	27.2	(28)	27.2	(28)		
Race							0.06	.80
Hispanic or Latino	13.6	(14)	7.8	(8)	5.8	(6)		
Not Hispanic or Latino	86.4	(89)	42.7	(44)	43.7	(45)		
Ethnicity							2.39	.66
Asian	1.0	(1)	1.0	(1)	0	(0)		
Black or African American	45.6	(47)	21.4	(22)	24.3	(25)		
Hispanic or Latino	13.6	(14)	7.8	(8)	5.8	(6)		
White	34.0	(35)	18.4	(19)	15.5	(16)		
Two or more ethnicities	5.8	(6)	1.9	(2)	3.9	(4)		
Special Education Eligibility							0.75	.39
General Education	94.2	(97)	48.5	(50)	45.6	(47)		
Special Education	5.8	(6)	1.9	(2)	3.9	(4)		
	M	*SD*	*M*	*SD*	*M*	*SD*	F	*p*
Age	8.10	0.05	8.11	0.05	8.10	0.04	1.37	.24

Table 5

Students' Average Scores on Initial Measures of Writing Performance

	Practice-Only		Performance Feedback			
Measure	*M*	*(SD)*	*M*	*(SD)*	*df*	*t*
CBM-WE[a]	19.37	(10.62)	14.53	(8.94)	101	2.50*
TOWL[b]	93.59	(10.71)	98.71	(8.97)	67	2.14*
Paragraph Copying Task[c]	21.06	(6.55)	21.31	(5.89)	98	0.05

Notes. [a]Curriculum-Based Measurement in Written Expression, as measured by number of correct writing sequences. [b]Standard score on the Test of Written Language- Third Edition with M =100 and SD = 15. [c]Measured by number of correctly copied words.

*$p < .05$.

Table 6

Descriptive Statistics and Correlations between Initial Measures of Writing Performance

	Total Sample					
Measure	*M*	*(SD)*	1	2	3	4
1. Paragraph Copying	21.18	(6.20)	--			
2. TOWL-3	96.71	(9.94)	.23	--		
3. CWS[a]	16.97	(10.07)	.49**	.21	--	
4. TWW[a]	20.13	(10.07)	.50**	.04	.94**	--

[a]Metric obtained from the Curriculum-Based Measurement in Written Expression probe.

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

Table 7

Teachers' Mean Scores on the Writing Orientation Scale

Factor	*M*	*SD*
Correct Writing[a]	3.03	1.25
Explicit Instruction[b]	4.67	1.13
Natural Learning[c]	4.29	0.95

Notes. $N = 6$. Answers were based on a Likert-type scale with 1 = strongly disagree and 6 = strongly agree. Factor scores were obtained by computing the average score of each item within that factor.

[a]Items 1, 5, 7, 11, 12. [b]Items 8, 9, 10, 13. [c]Items 2, 3, 4, 6.

Table 8

Ratings of Teachers' Instructional Practices

Item	Never	Several Times a Year	Monthly	Weekly	Several Times a Week	Daily	Several Times a Day
How often are specific writing strategies modeled to your students?	0	0	33%	67%	0	0	0
How often do you re-teach writing skills and strategies?	0	17%	33%	17%	17%	17%	0
How often do you conference with students about their writing?	0	33%	33%	33%	0	0	0
How often do students share their writing with their peers?	0	50%	0	17%	33%	0	0
How often do students help each other with their own writing?	0	33%	17%	17%	33%	0	0
How often do students select their own writing topics?	17%	0	50%	0	33%	0	0
How often do students use invented spelling in their writing?	0	0	0	0	33%	67%	0
How often do you specifically teach handwriting skills?	17%	17%	33%	0	17%	17%	0
How often do you specifically teach spelling skills?	0	17%	0	17%	50%	17%	0
How often do you specifically teach grammar skills?	0	17%	17%	33%	0	33%	0
How often do you specifically teach planning and revising strategies in writing?	0	33%	17%	17%	33%	0	0

Notes. N = 6.

Table 9

Adjusted Means, Standard Deviations, and ANCOVA Results for Stimulus Generalization Measure

	Performance Feedback				Practice-Only					
	Baseline		Post-Intervention		Baseline		Post-Intervention		ANCOVA	
	M	SD	M	SD	M	SD	M	SD	F (1, 87)	Partial η^2
Correct Writing Sequences	21.25	10.95	24.36	12.40	23.17	11.47	20.09	10.37	5.62*	.062

*p = .02.

Table 10

Adjusted Means, Standard Deviations, and ANCOVA Results for Stimulus-Response Generalization Measure

	Performance Feedback				Practice-Only					
	Baseline		Post-Intervention		Baseline		Post-Intervention		ANCOVA	
	M	SD	M	SD	M	SD	M	SD	F (1, 87)	Partial η^2
Correct writing sequences	73.81	37.88	69.87	40.66	61.96	35.16	66.33	31.64	.31*	.004

*p = .58.

Table 11

Mean Correct Writing Sequences, Standard Deviations, Percent Change, and T-test Results for Maintenance Probes

	Performance Feedback			Practice-Only				
	M	*SD*	Mean Percent Change	*M*	*SD*	Mean Percent Change	*t* (*df*)	*d*
Final Intervention Session	31.62	16.13	--	23.59	10.50	--	-- --	
2-Week Maintenance	27.82	15.38	1.58	28.0	11.86	31.80	-2.50 (80)*	-.55
4-Week Maintenance	25.51	12.84	-10.56	26.55	10.84	42.93	-3.32(51.08)**	-.93
6-Week Maintenance	28.26	13.87	-5.25	24.33	12.79	21.44	-1.40(43.85)	-.42

Notes. Mean percent change was calculated by averaging the percent change score for each individual.
*$p < .05$. **$p < .01$.

Table 12

Correlations between Final Intervention Probe and Maintenance Probes

	Performance Feedback				Practice-Only			
Measure	1	2	3	4	1	2	3	4
1. Final Intervention Probe	--				--			
2. 2-Week Maintenance Probe	.82***	--			.57***	--		
3. 4-Week Maintenance Probe	.69***	.79***	--		.47**	.57***	--	
4. 6-Week Maintenance Probe	.76***	.72***	.75***	--	.35*	.41*	.60***	--

*$p < .05$. **$p < .01$. **$p < .001$.

Table 13

Changes in Instructional Level

	Feedback Condition			
	Baseline		Post-Assessment	
Instructional Level	%	(*n*)	%	(*n*)
Frustrational	94.1	(48)	34.0	(16)
Instructional	2.0	(1)	14.9	(7)
Mastery	3.9	(2)	51.1	(24)
	Practice-Only Condition			
	Baseline		Post-Assessment	
Instructional Level	%	(*n*)	%	(*n*)
Frustrational	90.4	(47)	78.3	(36)
Instructional	3.8	(2)	8.7	(4)
Mastery	5.8	(3)	13.0	(6)
	Total Sample			
	Baseline		Post-Assessment	
Instructional Level	%	(*n*)	%	(*n*)
Frustrational	92.2	(95)	55.9	(52)
Instructional	2.9	(3)	11.8	(11)
Mastery	4.9	(5)	32.3	(30)

Note. Frustrational Level = 36 or fewer words written per 3 minutes. Instructional Level = 37 to 40 words written per 3 minutes. Mastery Level = 41 or more words written per 3 minutes.

Table 14

Students' Intervention Acceptability Ratings

	Total Sample[b]		Feedback[c]		Practice-Only[d]	
Procedures associated with CBM-WE	*M*	(*SD*)	*M*	(*SD*)	*M*	(*SD*)
How much do you like writing stories with us each week?	4.11	(1.27)	4.17	(1.10)	4.04	(1.43)
How much do you like being told what to write about?	3.27	(1.56)	3.33	(1.52)	3.22	(1.60)
Were there times when you didn't want to write stories with us?[a]	4.03	(1.30)	4.0	(1.37)	4.07	(1.25)
Were there any times when you wished you could work more on writing stories with us?	3.82	(1.51)	3.89	(1.44)	3.74	(1.58)
Procedures associated with performance feedback	*M*	(*SD*)	*M*	(*SD*)	*M*	(*SD*)
How much do you like being told how many words you wrote?			4.26	(1.25)		
How much do you think it helps you when you were told how many words you wrote?			4.21	(1.17)		
Procedures associated with practice	*M*	(*SD*)	*M*	(*SD*)	*M*	(*SD*)
Do you think your writing has improved?	4.37	(1.01)	4.35	(0.92)	4.39	(1.11)
Do you think your writing has gotten worse? [a]	4.74	(0.64)	4.87	(0.40)	4.61	(0.80)
Overall acceptability			4.14	(1.25)	3.91	(1.39)

Notes. Answers were based on a Likert-type scale with 1 = not at all, and 5 = very, very much. CBM-WE = Curriculum-Based Measurement in Written Expression.

[a]Item reversed scored so that higher numbers represent higher acceptability. [b]n = 92. [c]n = 46. [d]n = 46.

Table 15

Teachers' Intervention Acceptability Ratings

Item	*M*	*(SD)*
This would be an acceptable intervention for students' writing difficulties.	5.0	(0.63)
Most teachers would find this intervention appropriate for writing difficulties in addition to the one described.	4.50	(0.84)
This intervention should prove effective in changing students' writing difficulties.	4.33	(0.52)
I would suggest the use of this intervention to other teachers.	4.67	(0.82)
The students' writing difficulties are severe enough to warrant the use of this intervention.	5.50	(0.55)
Most teachers would find this intervention suitable for the writing difficulties described.	4.33	(0.52)
I would be willing to use this intervention in my classroom.	4.83	(0.75)
This intervention would not result in negative side effects for the students.	5.17	(0.41)
This intervention would be appropriate for a variety of students.	5.0	(0.63)
This intervention is consistent with those I have used in school.	3.50	(2.07)
The intervention is a fair way to handle the students' writing difficulties.	4.67	(0.82)
This intervention is reasonable for the writing difficulties described.	4.67	(0.82)
I like the procedures used in this intervention.	5.17	(0.75)
This intervention is a good way to handle the students' writing difficulties.	4.50	(0.84)
Overall, this intervention would be beneficial for the students.	4.67	(0.82)
Overall acceptability	4.70	(0.92)

Notes. $N = 6$. Answers were based on a Likert-type scale with 1 = strongly disagree, and 6 = strongly agree.

Figure 1. Hayes and Flower (1980) Model of Writing

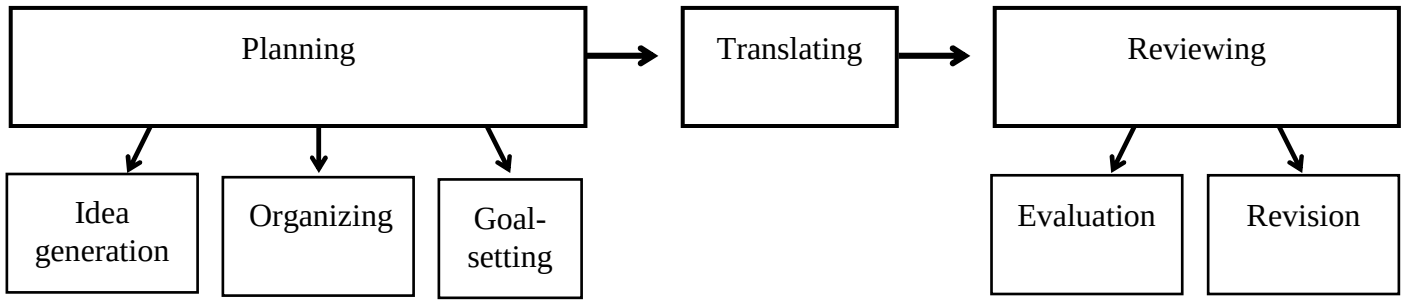


Figure 2. Berninger and colleagues' (1992) Component Processes of Writing

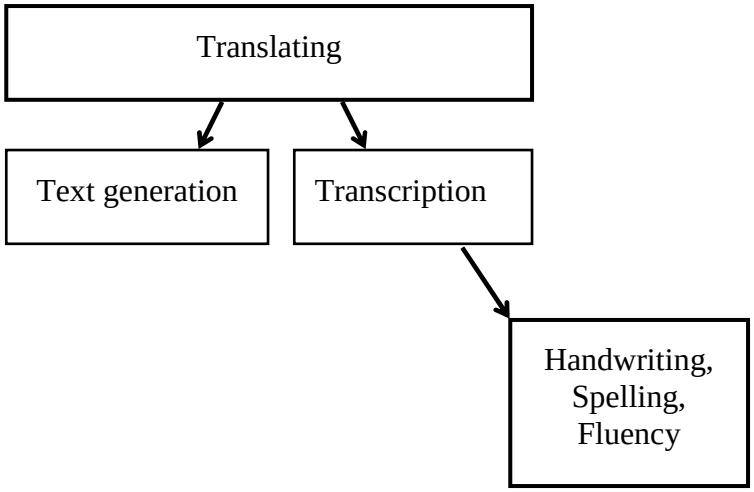
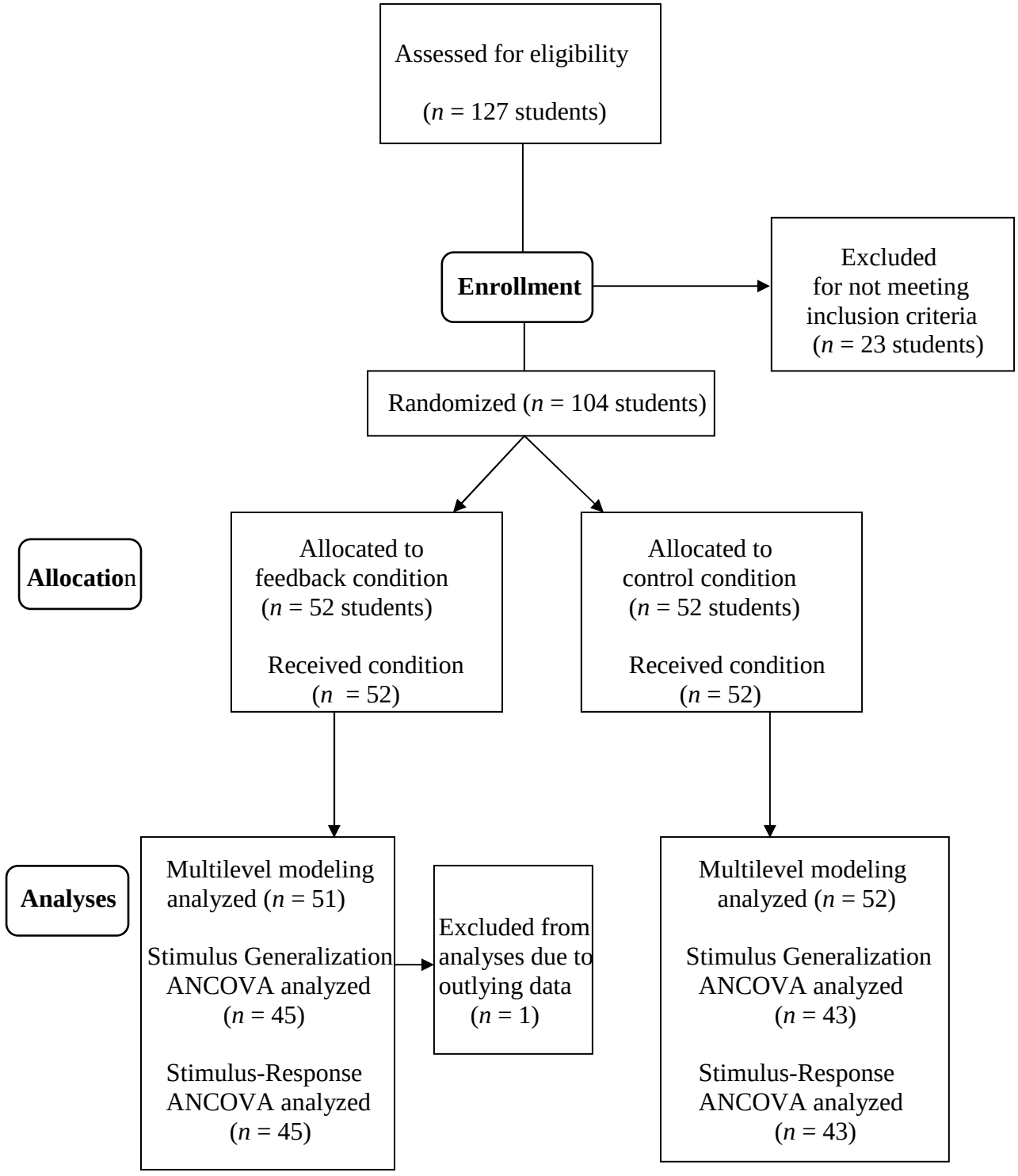


Figure 3. Participant flow chart following Consolidated Standards of Reporting Trials guidelines



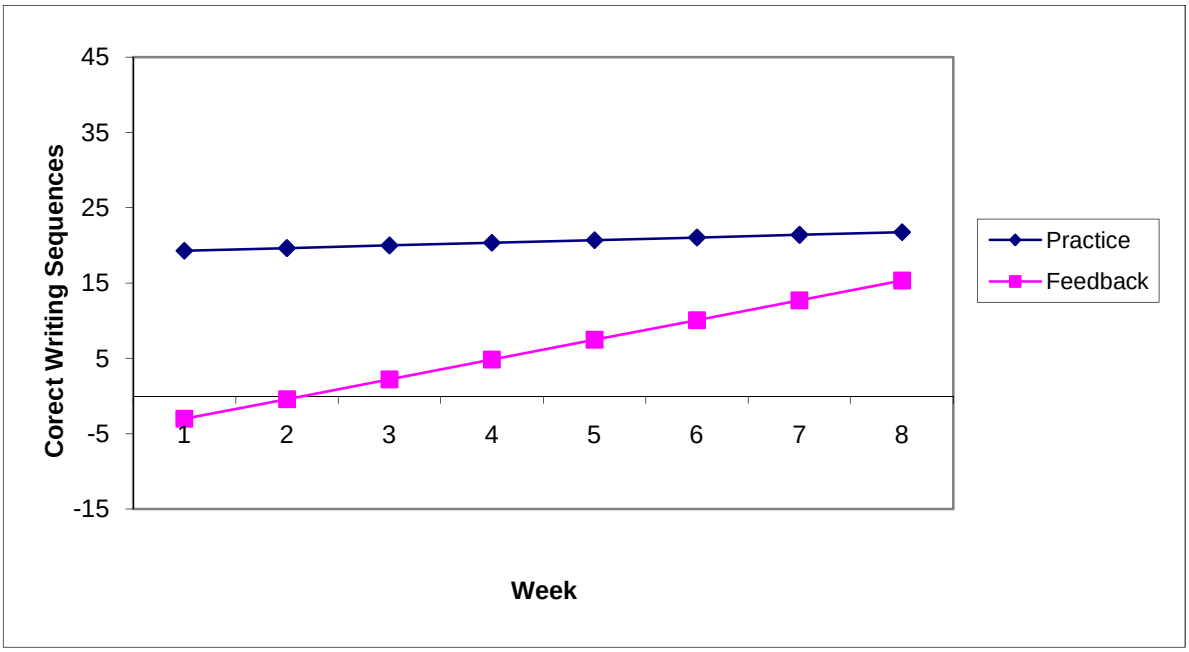


Figure 4. Growth trajectories by condition for students who wrote at frustrational level at baseline, reflecting students' average gains of correct writing sequences. Week 1 = baseline estimate computed through multilevel modeling procedure. Week 2 = true baseline. Weeks 3 – 8 = intervention phase.

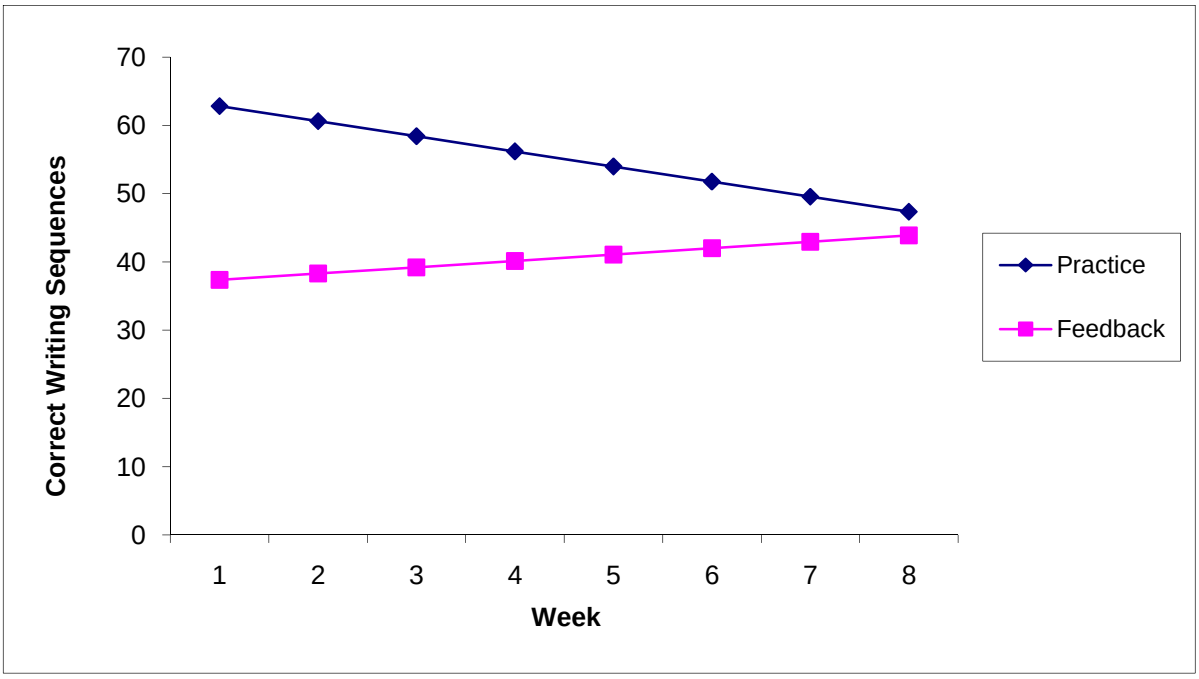
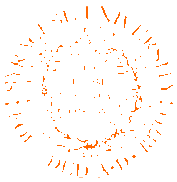


Figure 5. Growth trajectories by condition for students who wrote at mastery level at baseline, reflecting students' average gains of correct writing sequences.

Appendix A

Parent Informational Letter



SYRACUSE UNIVERSITY

COLLEGE OF ARTS AND SCIENCES
Department of Psychology

PARENT INFORMATIONAL LETTER

Treatment Research in Academic Competence
Examining Elementary-Aged Children's Written Expression Skills

Principal Investigator: Ms. Bridget Hier
Dept. of Psychology, Syracuse University
Phone: (315) 247-1978
Co-Principal Investigator: Dr. Tanya Eckert
Dept. of Psychology, Syracuse University
Phone: (315) 443-3141

Dear Parent or Guardian,

My name is Bridget Hier and I am a graduate student in the Department of Psychology at Syracuse University. I am working on a research study in your child's school in an attempt to better understand and improve children's writing skills. I am trying to see how much children's writing skills improve over time.

The purpose of this study is to determine how much children's academic skills change over time when given weekly feedback with writing practice. Beginning in February, myself and other students from Syracuse University will be working with your child's classroom for 15 minutes per week. During those 15 minutes, students will be told how they are doing in writing in addition to practicing writing.

If for any reason you do not want your child to participate in this study, please call me at 315-247-1978. Your decision will **NOT** affect your child's grades or your child's educational program.

Thank you!

Appendix B

Student Assent

Important Question

I would like to work with you each week for the next couple of months. We will be working on writing stories. Your parent has said that it is okay that I work with you. However, I want to make sure that it is okay with you. If you change your mind it is okay to stop working with me at any time.

Would it be okay if I worked with you on writing?

Yes ***No*** ***I don't know***

Name:_____

Appendix C

Writing Packet: Page 1, Identification Information

Syracuse University

2011-2012 Writing Project

_____Elementary School

3rd grade

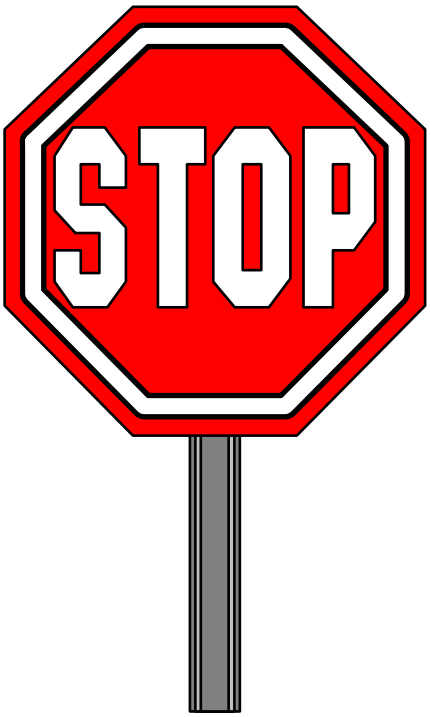
Name: _____

Classroom:

Probe # _____

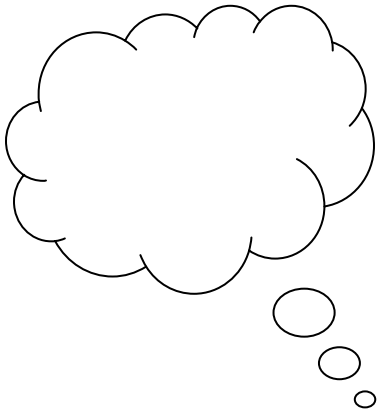
Appendix D

Writing Packet: Page 2, Stop Sign

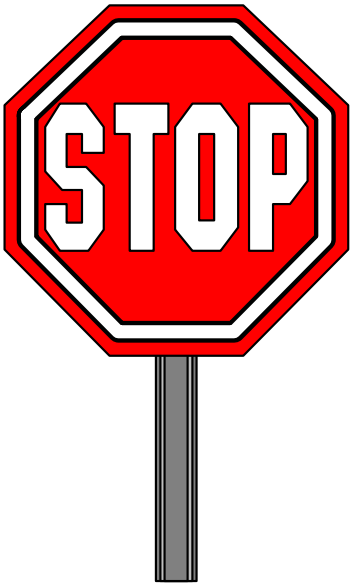


Appendix E

Writing Packet: Story Starter Page with Stop Sign



I was talking to my friends when all of a sudden . . .



Appendix F

Writing Packet: Story Starter with Writing Lines

I was talking to my friends when all of a sudden _____

Keep going 

Appendix G

Stimulus-Response Generalization Task

WRITTEN COMPOSITION

Write a composition about how gym class is the same as and different from art class.

The information in the box below will help you remember what you should think about when you write your composition.

REMEMBER TO –

- write about how gym class is the same as and different from art class
- make sure that every sentence you write helps the reader understand your composition
- include enough details to help the reader clearly understand what you are saying
- use correct spelling, capitalization, punctuation, grammar, and sentences

USE THIS PREWRITING PLANNING PAGE TO
PLAN YOUR COMPOSITION

Answer Document

Answer Document

Answer Document

Appendix H:

Procedural Script for Stimulus-Response Generalization Task

Directions: Please fill out each area detailed below. Please make sure that the identifying information (box 1) is complete before you submit the form.

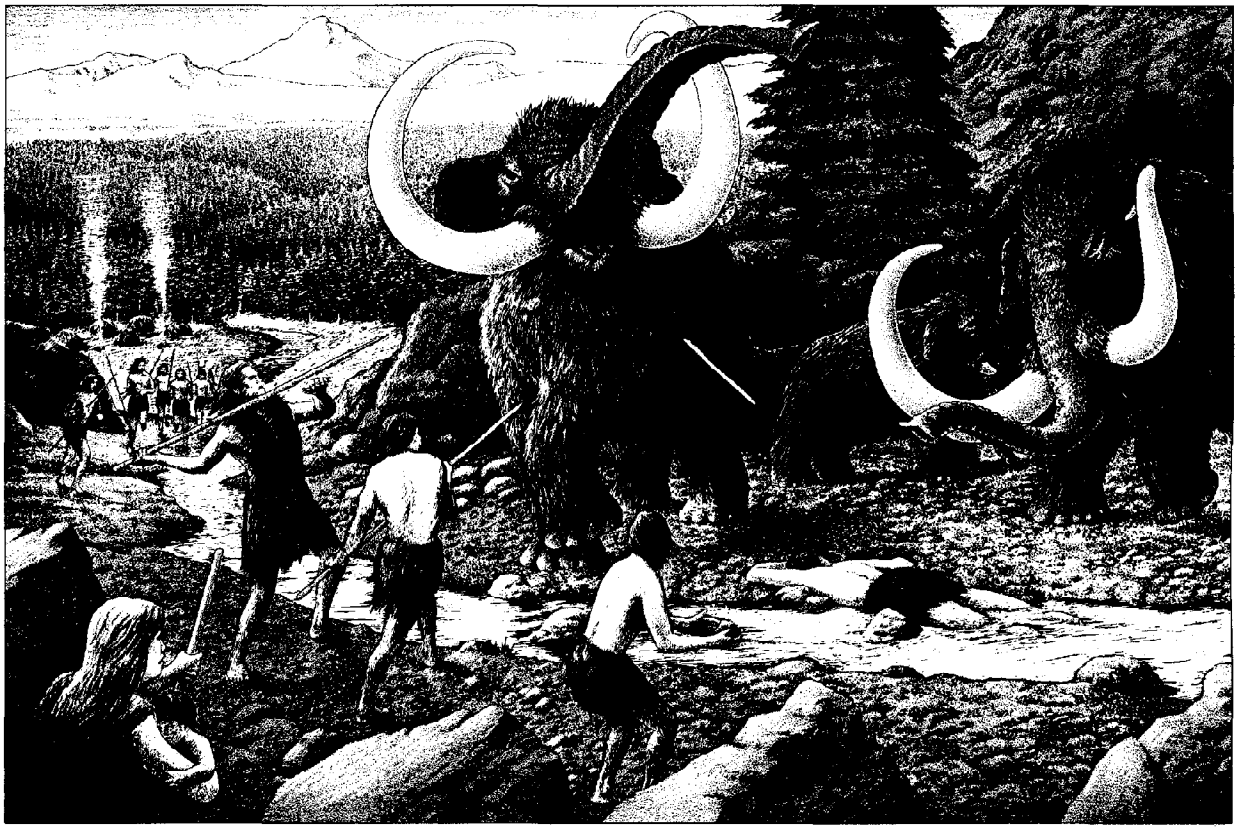
I.	Identifying Information		
Name of primary research assistant (general classroom teacher):			
Name of secondary research assistant:		or N/A	
School/Classroom:			
Notes:			
II.	Data Collection – Material Preparation	Circle	
a.	Five (5) sharpened pencils	Yes	No
b.	Assessment packets	Yes	No
c.	Experimenter's copy of packet	Yes	No
d.	Two (2) stopwatches	Yes	No
Notes:			
III.	Data Collection Procedures [Please check [✓] each box as you complete each step]✓		
1.	State to the students: "*If you haven't already done so, please clear everything off of your desk except for a pencil. I will be passing around packets. When you get yours, please keep it closed and quietly wait for my instructions.*"		
2	Teacher should distribute the packets.		
3.	After all of the packets have been distributed, state to the students: "*Please turn to the second page of your packet, which looks like this."* [Hold up the correct page for students to see.]		
4.	Teacher should make sure all the students are on the correct page.		
5.	State to the students: "*Look at the writing prompt on this page.*" [Teacher should hold up the packet and point to the writing prompt.]"*It is followed by a planning page on the next page. You may use that blank page to plan your composition. You may make notes to help you decide what you want to write. In addition, you may write an outline to help you arrange your ideas in an order that makes sense. Remember that the more planning you do, the clearer and more complete your composition is likely to be. For the next few minutes, take time to plan your composition. Do not write your*		

	composition yet; just use the planning page to make notes if you would like to."	
6.	Teacher should start the stopwatch and time students for 3 minutes.	
7.	During this time, teacher should monitor students to ensure they are following directions from step #5, and ensure they are <u>not</u> writing on the composition pages with lines. For any student writing on the lined composition page, redirect students by flipping to the planning page and saying, *"Please just plan right now – I'll tell you when to begin writing the composition."* Do not prompt any students who are not writing on the planning page.	
8.	At the end of 3 minutes, state to the students: *"Please turn to the next page with lines on it."*	
9.	Teacher should monitor to ensure all students have turned to the correct page and have not started writing yet.	
10.	State to the students: *"When I tell you to begin, start writing your composition on this page with lines. If you need more paper, turn to the next page. Your composition does not have to completely fill these two lined pages. Are there any questions?"*	
11.	Answer questions so that all students understand what they are supposed to do.	
12.	State to the students: *"Okay, you may begin writing now."*	
13.	Teacher should discretely start the stopwatch and time the students for 10 minutes.	
14.	During this time, the teacher should monitor to ensure they are following directions from step #10. Do not prompt students who stop writing early.	
15.	At the end of 10 minutes, state to the students: *"Please stop writing and close your packet."*	
16.	Monitor to ensure all students have stopped writing.	
17.	Collect all student packets.	

Appendix I

Test of Written Language-Third Edition Spontaneous Writing prompt

STORY



Appendix J

Paragraph Copying Task

A little boy lived with his father in a large forest. Every day the father went out to cut wood. One day the boy was walking through the woods with a basket of lunch for his father. Suddenly he met a huge bear. The boy was frightened, but he threw a piece of bread and jelly to the bear.

Appendix K

Handwriting Proficiency Screening Measure

Please wait for our directions.

Please print each letter that is spoken.

1.

2.

3.

4.

5.

6.

7.

8.

9.

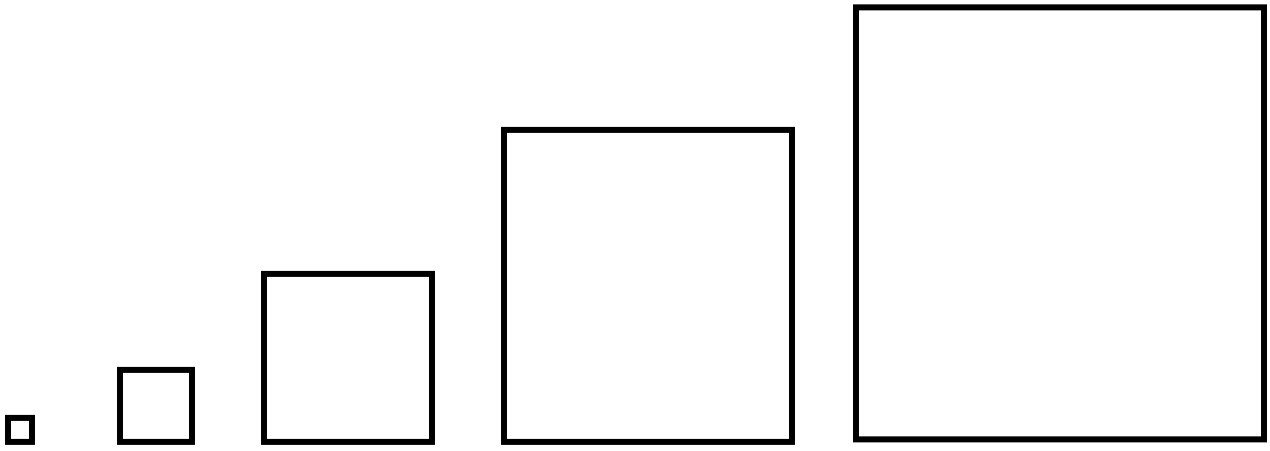
10.

Appendix L

Kids Intervention Profile

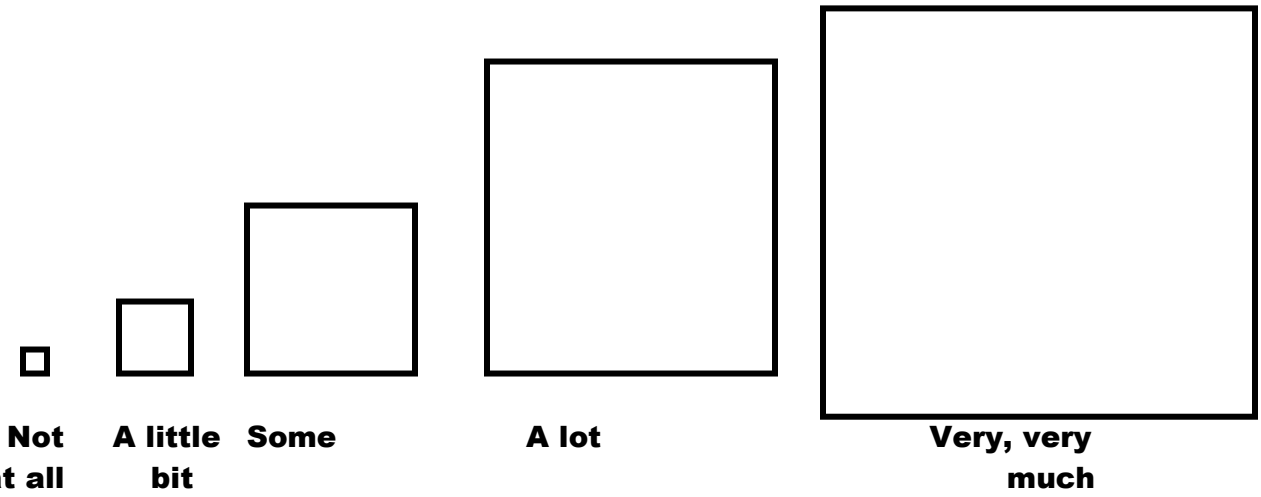
Question #1

How much do you like writing stories with us each week?



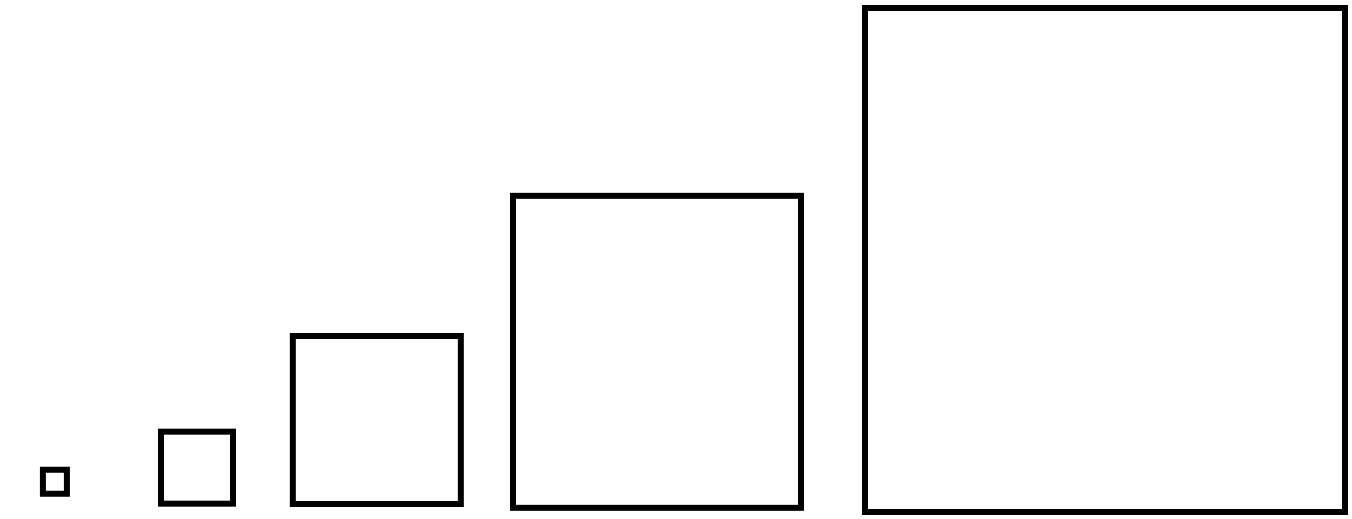
Question #2

How much do you like being told what to write about?



Question #3

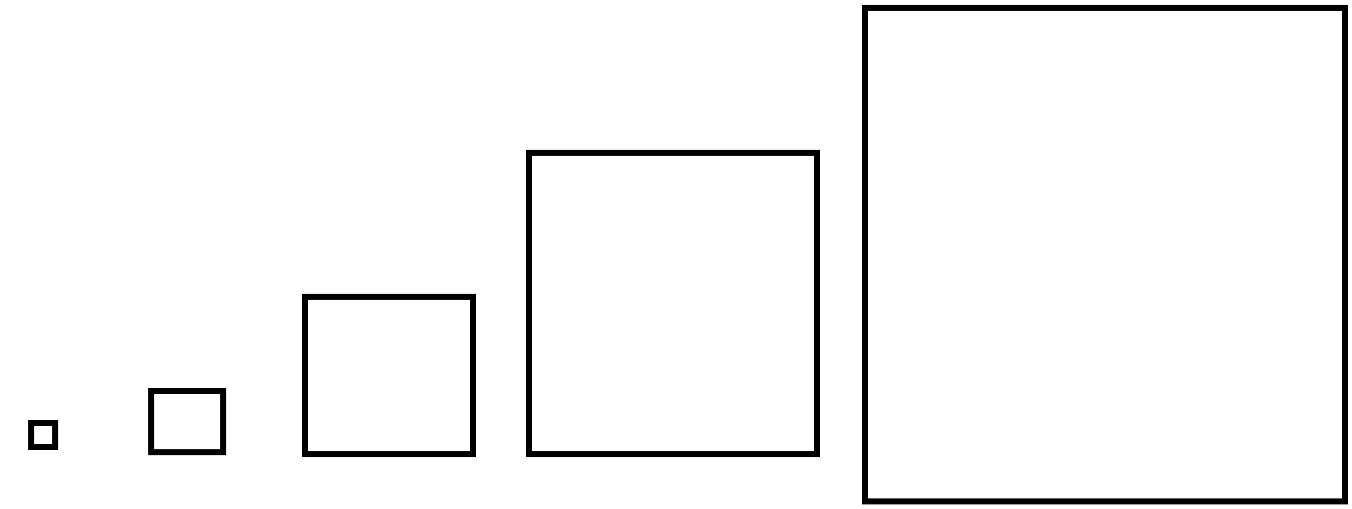
Were there times when you didn't want to write stories with us?



Never | **A couple times** | **Sometimes** | **A lot of times** | **Many, many times**

Question #4

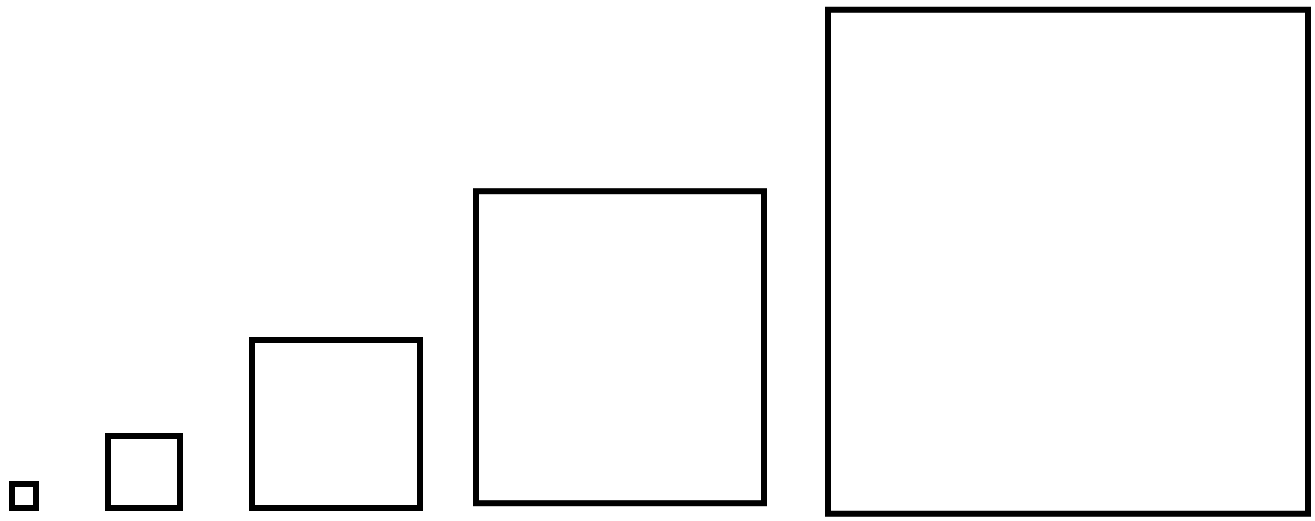
Were there any times when you wished you could work more on writing stories with us?



Never | **A couple times** | **Sometimes** | **A lot of times** | **Many, many times**

Question #5

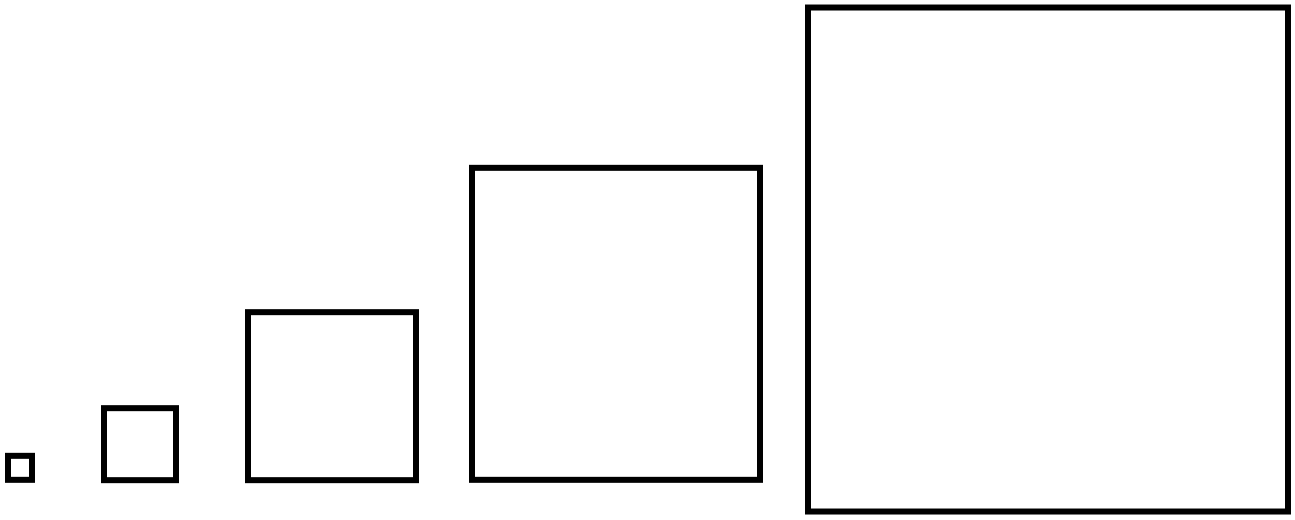
How much do you like being told how many words you wrote?



Not at all **A little bit** **Some** **A lot** **Very, very much**

Question #6

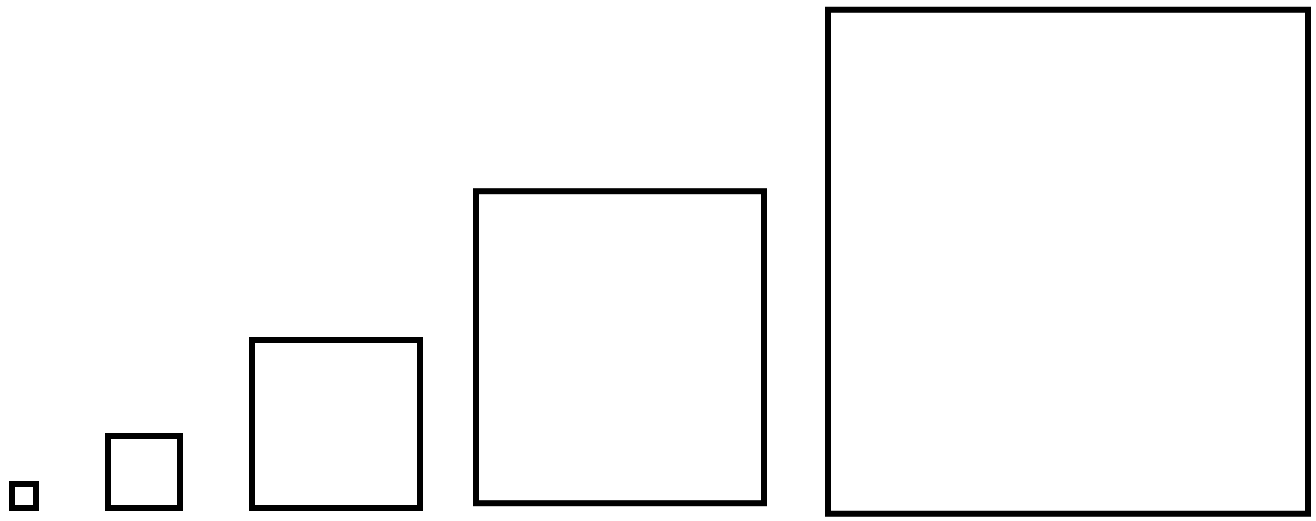
How much do you think it helps you when you weretold how many words you wrote?



Not at all **A little bit** **Some** **A lot** **Very, very much**

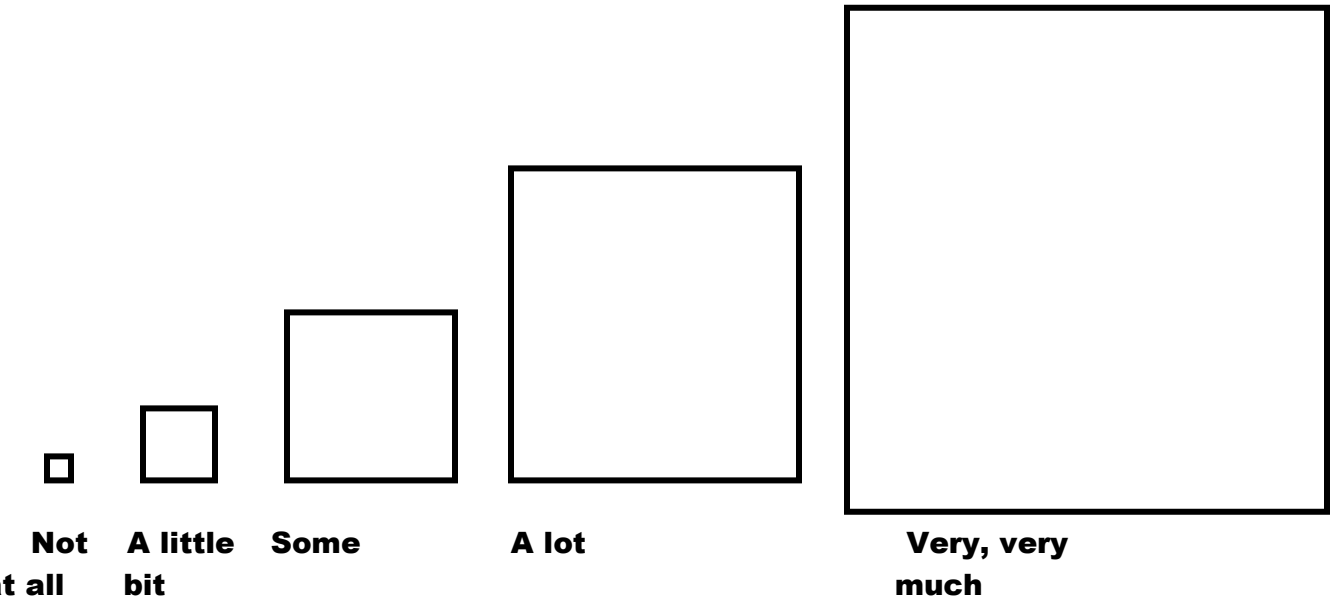
Question #7

Do you think your writing has improved?



Question #8

Do you think your writing has gotten worse?



Appendix M

Intervention Rating Profile – 15 (IRP-15) – Teacher Version

Teacher's Name: _____ Date: _____

The purpose of this questionnaire is to get information that will help in the selection of treatments for children. Please circle the number which best describes your agreement or disagreement with each statement.

	Strongly Disagree	Disagree	Slightly Disagree	Slightly Agree	Agree	Strongly Agree
1. This would be an acceptable intervention for students' writing difficulties.	1	2	3	4	5	6
2. Most teachers would find this intervention appropriate for writing difficulties in addition to the one described.	1	2	3	4	5	6
3. This intervention should prove effective in changing students' writing difficulties.	1	2	3	4	5	6
4. I would suggest the use of this intervention to other teachers.	1	2	3	4	5	6
5. The students' writing difficulties are severe enough to warrant the use of this intervention.	1	2	3	4	5	6
6. Most teachers would find this intervention suitable for the writing difficulties described.	1	2	3	4	5	6
7. I would be willing to use this intervention in my classroom.	1	2	3	4	5	6
8. This intervention would not result in negative side effects for the students.	1	2	3	4	5	6

9. This intervention would be appropriate for a variety of students.	1	2	3	4	5	6
10. This intervention is consistent with those I have used in school.	1	2	3	4	5	6
11. The intervention is a fair way to handle the students' writing difficulties.	1	2	3	4	5	6
12. This intervention is reasonable for the writing difficulties described.	1	2	3	4	5	6
13. I like the procedures used in this intervention.	1	2	3	4	5	6
14. This intervention is a good way to handle the students' writing difficulties.	1	2	3	4	5	6
15. Overall, this intervention would be beneficial for the students.	1	2	3	4	5	6

Appendix N

Teacher Questionnaire

Teacher's name: _____ Date: _____

Directions: Please answer the following questions so we may know more about your professional and educational experiences and credentials.

1) Total number of years of teaching: _____ years
2) Total number of years at current school: _____ years
3) Teaching degree(s): _____
4) Additional certification(s): _____

Classroom Events

We are also interested in some events that routinely occur in elementary classrooms. These events can make teaching more challenging. Please read each item and circle how often it happens in your classroom (rarely, sometimes, a lot, or constantly) and then circle how much of a `challenge' you feel that it has been for you during the past 3 months.

Event	**How often it happens**				**Challenge (low to high)**
Cleaning up classroom messes	*Rarely*	*Sometimes*	*A lot*	*Constantly*	1 2 3 4 5
Being nagged and complained to	*Rarely*	*Sometimes*	*A lot*	*Constantly*	1 2 3 4 5
Students will not follow directions	*Rarely*	*Sometimes*	*A lot*	*Constantly*	1 2 3 4 5
Student arguments requiring a `referee'	*Rarely*	*Sometimes*	*A lot*	*Constantly*	1 2 3 4 5
Students demand constant attention	*Rarely*	*Sometimes*	*A lot*	*Constantly*	1 2 3 4 5
Students resist completing school assignments	*Rarely*	*Sometimes*	*A lot*	*Constantly*	1 2 3 4 5
Students interrupt adult conversations	*Rarely*	*Sometimes*	*A lot*	*Constantly*	1 2 3 4 5
Having to change lesson plans due to unprecedented student learning or behavioral needs	*Rarely*	*Sometimes*	*A lot*	*Constantly*	1 2 3 4 5
Difficulties in managing students during classroom transitions	*Rarely*	*Sometimes*	*A lot*	*Constantly*	1 2 3 4 5
Students arrive to school late	*Rarely*	*Sometimes*	*A lot*	*Constantly*	1 2 3 4 5
Difficulties in getting students ready for school dismissal	*Rarely*	*Sometimes*	*A lot*	*Constantly*	1 2 3 4 5
Students are removed from the classroom due to disciplinary issues	*Rarely*	*Sometimes*	*A lot*	*Constantly*	1 2 3 4 5
Students' parents frequently make complaints	*Rarely*	*Sometimes*	*A lot*	*Constantly*	1 2 3 4 5

Writing Instruction

The purpose of our work is to examine effective writing strategies for students in elementary school. It would be helpful if you could identify any specific writing curricula or programs that you use to develop your writing lesson plans:

_____

_____

Teaching Philosophy in Writing

In addition, we are interested in learning more about your teaching philosophy regarding written expression. Please answer the following questions:

#	Item	Strongly Disagree	Disagree	Slightly Disagree	Slightly Agree	Agree	Strongly Agree
1	A good way to begin writing instruction is to have children copy good models of each particular type of writing.	1	2	3	4	5	6
2	Instead of regular grammar lessons, it is best to teach grammar when a specific need for it emerges in a child's writing.	1	2	3	4	5	6
3	Students need to meet frequently in small groups to react and critique each other's writing.	1	2	3	4	5	6
4	The act of composing is more important than the written work children produce.	1	2	3	4	5	6
5	Before children begin a writing task, teachers should remind them t use correct spelling.	1	2	3	4	5	6
6	With practice writing and responding to written messages, children will gradually learn the conventions of adult writing.	1	2	3	4	5	6
7	Being able to label words according to grammatical function (e.g., nouns and verbs) is useful in proficient writing.	1	2	3	4	5	6
8	It is important for children to study words in order to learn their spell	1	2	3	4	5	6
9	Formal instruction in writing is necessary to insure adequate development of all the skills used in writing.	1	2	3	4	5	6
10	Children need to practice writing letters to learn how to form them correctly.	1	2	3	4	5	6
11	Teachers should aim at producing writers who can write good compositions in one draft.	1	2	3	4	5	6
12	Before they begin a writing task, children who speak a non-standard dialect of English should be reminded to use correct English.	1	2	3	4	5	6
13	It is important to teach children strategies for planning and revising.	1	2	3	4	5	6

Instructional Practices in Writing

Next, we are interested in learning more about your instructional practices in writing. Please answer the following questions:

#	Item	Never	Several times a year	Monthly	Weekly	Several times a week	Daily	Several times a day
1	How often are specific writing strategies modeled to your students?	1	2	3	4	5	6	7
2	How often do you re-teach writing skills and strategies?	1	2	3	4	5	6	7
3	How often do you conference with students about their writing?	1	2	3	4	5	6	7
4	How often do students share their writing with their peers?	1	2	3	4	5	6	7
5	How often do students help each other with their writing?	1	2	3	4	5	6	7
6	How often do students select their own writing topics?	1	2	3	4	5	6	7
7	How often do students use invented spelling in their Writing?	1	2	3	4	5	6	7
8	How often do you specifically teach handwriting skills?	1	2	3	4	5	6	7
9	How often do you specifically teach spelling skills?	1	2	3	4	5	6	7
10	How often do you specifically teach grammar skills?	1	2	3	4	5	6	7
11	How often do you specifically teach planning and revising strategies in writing?	1	2	3	4	5	6	7

Instructional Time in Writing

Finally, we are interested in learning how much instructional time is allocated for different writing activities. Please estimate **how many minutes per week** students in your classroom are engaged in:

(1) Handwriting practice: _____ minutes
(2) Spelling practice: _____ minutes
(3) Composition writing: _____ minutes

Appendix O

Procedural Script for Individualized Performance Feedback Condition

Directions: Please fill out each area detailed below. Please make sure that the identifying information (box 1) is complete before you submit the form.

I.	Identifying Information		
Name of primary research assistant:			
Name of secondary research assistant:		or N/A	
School/Classroom:			
Date:			
Notes:			
II.	Data Collection – Material Preparation	Circle	
a.	Five (5) sharpened pencils	Yes	No
b.	Assessment packets	Yes	No
c.	Experimenter's copy of packet	Yes	No
d.	Two (2) stopwatches	Yes	No
e.	Insert names	Yes	No
Notes:			
III.	Data Collection Procedures [Please check [✓] each box as you complete each step]✓		
1.	State to the students: "*Hello. If you have not already done so, please clean off the top of your desk, except for a pencil. Please listen for your name as _____ and I hand out the packets. Raise your hand when we call your name.*"		
2	Both research assistants should distribute the packets. (This should be very quick and not take longer than 2-3 minutes.)		
3.	After all of the packets have been distributed, State to the students: "*When I call your group color, please line up at the door with your packet and your pencil*." *"The Green Group will be staying in this classroom to work with us.* *"The Blue Group will be will be going to _____'s classroom. Please line up now and show me how you walk quietly through the halls at* _____ .		
4.	As students from other classes enter the classroom, the research assistant should direct students to sit down at the nearest desk in a systematic fashion. Do not let students talk you into letting them sit next to friends. Once the desks fill up, place any remaining students at tables in the room.		

	The other research assistant should be standing outside the classroom holding up the green sheet of paper that says Green Group. The research assistant should assist students with quickly getting to the appropriate classroom.	
5.	Once you have confirmed that all the students from the other classrooms have arrived, state to the students: "*Welcome to the Green Group. Please turn to the red page of your packet that has stop sign in the middle of the page. Today I want you to write a story. Before we do that I want to tell you how you are doing with your writing skills. Last week we took all your stories back to SU and we counted all of the words that each of you wrote in your stories. Please turn to the next page of your packet. This page has a funnel with some numbers going into it at the top of the page.*"	
6.	The research assistant should scan the room to make sure all the students are on the correct page.	
7	State to the students: *"The box in the middle of the page* [The research assistant should point to the box.] *tells you how many words you wrote last week. Next to the box you will see an arrow.* *If the arrow is pointing up towards the sky, you wrote more words since the last time I worked with you.* *If the arrow is pointing down towards the floor, that means you wrote fewer words since the last time I worked with you.* *Every week when I work with you, I will tell you how you are doing with your writing."*	
8	The research assistant should monitor the students for questions.	
9.	State to the students: "*Now I want you to write another story. I am going to read a sentence to you first, and then I want you to write a story about what happens next. You will have some time to think about the story you will write and then you will have some time to write it."*	
10.	State to the students: "*Please turn to the next page of your packet. This page has a thinking dog at the top of the page.*"	
11.	State to the students: "*For the next minute think about writing a story that begins with this sentence* – I was talking to my friends when all of a sudden. . . *Remember, take time to plan your story. A well-written story usually has a*	

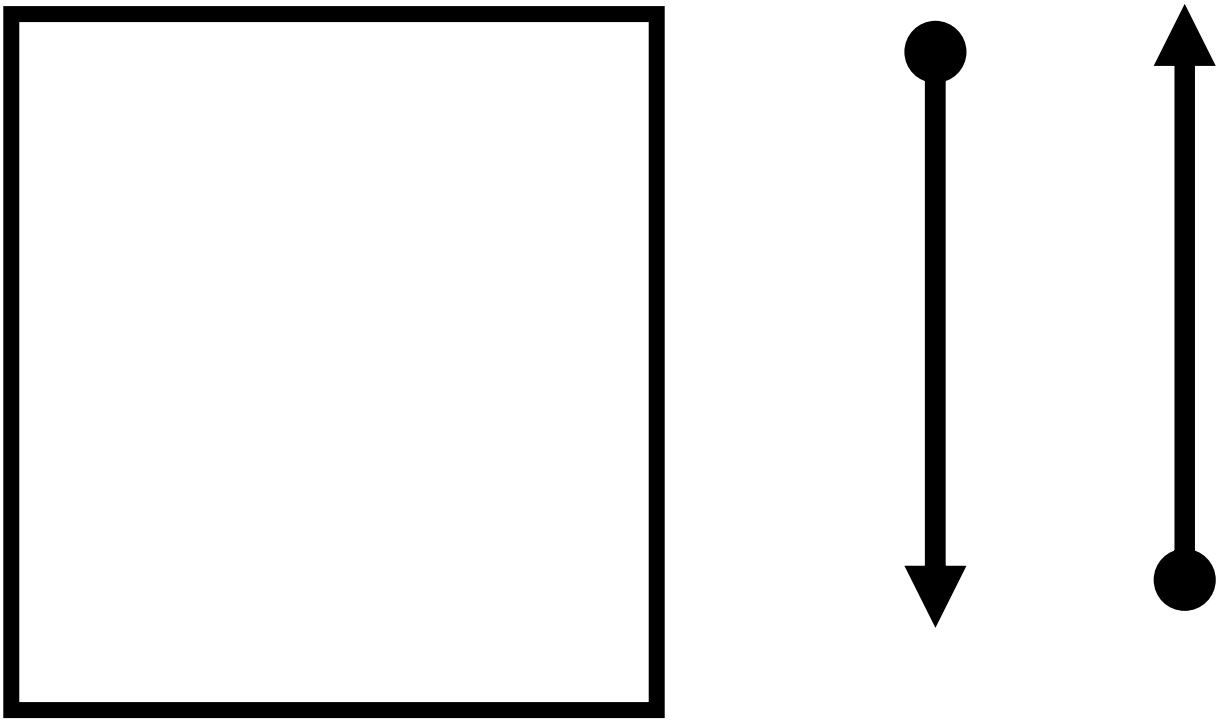
	beginning, a middle, and end. It also has characters that have names and perform certain actions. Use paragraphs to help organize your story. Correct punctuation and capitalization will make your story easier to read. *Please do not write the story. Just think of a story that begins with this sentence - –* <u>I was talking to my friends when all of a sudden. . .</u>"	
12.	The research assistant should begin the stopwatch and time the students for 1 minute.	
13.	At the end of 1 minute, state to the students: "*Okay, stop thinking, turn to the next page of your packet which has a bee holding a pencil, and raise your pencil in the air."*	
14.	State to the students: *"When I tell you to start, please begin writing your story. Remember, if you don't know how to spell a word, you should try your best and sound it out. It is important that you do your best work. If you fill up the first page, please turn to the next page and keep writing. Do not stop writing until I tell you to. Do your best work."*	
15.	State to the students: "*Okay, you can start writing.*" The research assistant should begin the stop watch and time the students for 3 minutes.	
16.	The research assistant should monitor the students during the 3-minute period and make sure students are following the directions stated in step #14. Also monitor the students to make sure that they are not re-copying the story starter. If a student is re-copying the starter, state to the student *"you do not need to copy the words that have been provided"*	
17.	After 1 minute, 30 seconds has elapsed, state to the students: *"You should be writing about* – <u>I was talking to my friends when all of a sudden</u> *"*	
18.	After 3 minutes has elapsed, state to the students: *"Please stop writing, put your pencils back in the air, and turn to the next page of your packet that has several boxes on it.*"	
19.	The research assistant should scan the room to make sure that all of the students have followed the directions. State to the students: *"Lastly, I want you to answer some questions about the story you just wrote. Look at question number 1. It says 'How much did you like writing about I was talking to my friends when all of a sudden . . .?' Make an 'X' through the box which tells how*	

	much you liked writing about I was talking to my friends when all of a sudden..."	
20.	Continue reading each of the remaining 4 questions.	
21	State to the students: *That is all of the writing that we are going to do today. All of you did a very nice job following my directions.* If the other classroom is not finished, state to the students: "*Please turn to the last page of the packet. This page has a word find on it. You may work quietly on this word find until the other classrooms are ready to switch."* When the other 2 classrooms are ready to switch, continue to step #23	
22.	State to the students: "Please hand in your packets. *Thank you for working with us today.*" If students complain about not finishing the word find, let them tear off the back page and tell them they may complete it at home.	
23.	The research assistant should collect all of the packets.	
24.	State to the students: *All of the students in _____'s classroom, please pick up your pencil and line up to the left side of the door.*	
25.	The research assistants should then assist the students in getting back to their classrooms quickly and quietly.	
Total number of steps completed:		

Appendix P

Feedback Page for Performance Feedback Condition

Here is how you are doing in writing:



Appendix Q

Procedural Script for Practice-Only Condition

Directions: Please fill out each area detailed below. Please make sure that the identifying information (box 1) is complete before you submit the form.

I.	Identifying Information		
Name of primary research assistant:			
Name of secondary research assistant:		or N/A	
School/Classroom:			
Date:			
Notes:			
II.	Data Collection – Material Preparation	Circle	
a.	Five (5) sharpened pencils	Yes	No
b.	Assessment packets	Yes	No
c.	Experimenter's copy of packet	Yes	No
d.	Two (2) stopwatches	Yes	No
e.	Insert names	Yes	No
Notes:			

III.	Data Collection Procedures [Please check [✓] each box as you complete each step]✓	
1.	State to the students: "*Hello. If you have not already done so, please clean off the top of your desk, except for a pencil. Please listen for your name as _____ and I hand out the packets. Raise your hand when we call your name.* "	
2	Both research assistants should distribute the packets. (This should be very quick and not take longer than 2-3 minutes.	
3.	After all of the packets have been distributed, State to the students: "*When I call your group color, please line up at the door with your packet and your pencil*." *"The Gold Group will be staying in this classroom to work with us.* *"The Green Group will be going to _____'s classroom. Please line up now and show me how you walk quietly through the halls at* _____ . *"The Blue Group will be will be going to _____'s classroom. Please line up now and show me how you walk quietly through the halls at* _____ .	

4.	As students from other classes enter the classroom, the research assistant should direct students to sit down at the nearest desk in a systematic fashion. Do not let students talk you into letting them sit next to friends. Once the desks fill up, place any remaining students at tables in the room. The other research assistant should be standing outside the classroom holding up the green sheet of paper that says Green Group. The research assistant should assist students with quickly getting to the appropriate classroom.	
5.	Once you have confirmed that all the students from the other classrooms have arrived, state to the students: "*Welcome to the Gold Group. Please turn to the red page of your packet that has stop sign in the middle of the page. Today I want you to write another short story. You will have some time to think about the story you will write and then you will have some time to write it."*	
6.	The research assistant should scan the room to make sure all the students are on the correct page.	
7.	State to the students: "*Please turn to the next page of your packet. This page has a thinking dog at the top of the page.*"	
8.	State to the students: "*For the next minute think about writing a story that begins with this sentence –* <u>I was talking to my friends when all of a sudden. . .</u> *Remember, take time to plan your story. A well-written story usually has a beginning, a middle, and end. It also has characters that have names and perform certain actions. Use paragraphs to help organize your story. Correct punctuation and capitalization will make your story easier to read.* *Please do not write the story. Just think of a story that begins with this sentence - –* <u>I was talking to my friends when all of sudden. . .</u>"	
9.	The research assistant should begin the stopwatch and time the students for 1 minute.	
10.	At the end of 1 minute, state to the students: "*Okay, stop thinking, turn to the next page of your packet which has a bee holding a pencil, and raise your pencil in the air."*	
11.	State to the students: *"When I tell you to start, please begin writing your story. Remember, if you don't know how to spell a word, you should try your best and sound it out. It is important that you do your best work. If you fill up the first page, please turn to the next page and keep writing. Do not stop writing until I tell you to. Do your best work."*	

12.	State to the students: "*Okay, you can start writing.*" The research assistant should begin the stop watch and time the students for 3 minutes.	
13.	The research assistant should monitor the students during the 3-minute period and make sure students are following the directions stated in step #38. Also monitor the students to make sure that they are not re-copying the story starter. If a student is re-copying the starter, state to the student *"you do not need to copy the words that have been provided"*	
14.	After 1 minute, 30 seconds has elapsed, state to the students: *"You should be writing about* - – <u>I was talking to my friends when all of sudden</u> *"*	
15.	After 3 minutes has elapsed, state to the students: *"Please stop writing, put your pencils back in the air, and turn to the next page of your packet that has several boxes on it.*"	
16.	The research assistant should scan the room to make sure that all of the students have followed the directions. State to the students: *"Lastly, I want you to answer some questions about the story you just wrote. Look at question number 1. It says 'I was talking to my friends when all of a sudden . . .?' Make an 'X' through the box which tells how much you liked writing about I was talking to my friends when all of a sudden ..."*	
17.	Continue reading each of the remaining 4 questions.	
18	State to the students: *That is all of the writing that we are going to do today. All of you did a very nice job following my directions.* If the other 2 classrooms are not finished, state to the students: "*Please turn to the last page of the packet. This page has a word find on it. You may work quietly on this word find until the other classrooms are ready to switch."* When the other 2 classrooms are ready to switch, continue to step #19	

19.	State to the students: "Please hand in your packets. *Thank you for working with us today*." If students complain about not finishing the word find, let them tear off the back page and tell them they may complete it at home.	
20.	The research assistant should collect all of the packets.	
21.	State to the students: *All of the students in _____'s classroom, please pick up your pencil and line up to the left side of the door. All of the students in _____'s classroom, please pick up your pencil and line up to the right side of the door.*	
22.	The research assistants should then assist the students in getting back to their classrooms quickly and quietly.	
Total number of steps completed:		

Appendix R

Scoring Manual

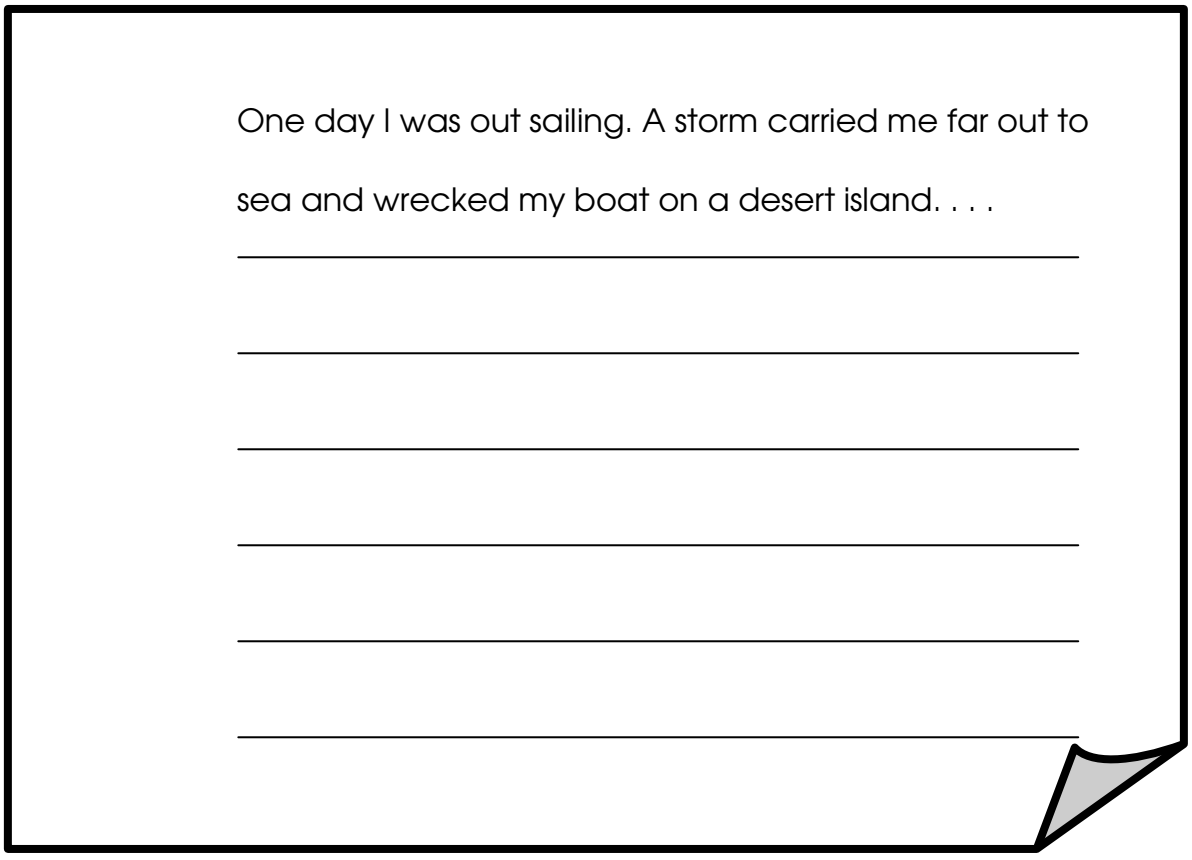
2009-2010 TRAC RESEARCH PROJECT

RA Training Manual: Administration and Scoring of Curriculum-Based Measurement in Written Expression Probes

Curriculum-Based Measurement - Introduction

Curriculum-Based Measurement (CBM) is an alternative measurement system that has been developed for assessing students' academic skills. CBM is designed to provide a reliable and direct estimate of students' skills. In addition, CBM is sensitive to measuring student growth over time. The measures collected are brief and repeatable, and generally consist of timed skill worksheets. These worksheets are often referred to as "**probes.**"

For the purposes of this project, we will be focusing on using CBM in the academic area of written expression (CBM-WE). CBM-WE emphasizes assessing basic writing fluency as the foundation upon which success in other aspects of writing are developed. To assess basic writing fluency, we will be providing students with a "**story starter**" and asking students to complete one story from the story starter during a relatively short period of time. The story stem appears at the top of a lined composition sheet. The student is instructed to think for **1 minute** about a possible story to be written from the story starter, then spends **3 minutes** writing the story. The examiner collects the writing sample for scoring. A sample CBM-WE probe appears below:



CBM-WE - Administration

Materials:

The following materials are needed for administering CBM-WE probes:

(1) the student's copy of the CBM-WE probe containing the story starter

(2) a stopwatch for the examiner

(3) a writing instrument (i.e., pencil) for the student

Administration:

The examiner distributes copies of the CBM-WE probes to all students being assessed. The examiner provides the following directions to the students:

I want you to write a story. I am going to read a sentence to you first, and then I want you to write a short story about what happens. You will have 1 minute to think about the story you will write and then have 3 minutes to write it. Do your best work. If you don't know how to spell a word, you should try your best to sound out the word. Are there any questions?

For the next minute, think about . . . (insert story starter). The examiner starts the stopwatch.

At the end of 1 minute, the examiner says, *Start writing.*

While the students are writing, the examiner and any other adults helping with the data collection circulate around the room. If students stop writing before the 3 minute timing period has ended, the adults encourage them to continue writing.

After 3 additional minutes, the examiner says, *Stop writing. Please put your pencils down.*

CBM-WE probes are collected for scoring.

Curriculum-Based Measurement - Scoring

There are several options when scoring CBM-WE probes. Student writing samples may be scored according to the:

(1) number of total words written (TWW)

(2) number of correctly spelled words (CSW)

(3) number of writing units placed in correct sequence – correct word sequences (CWS)

(4) incorrect writing sequences (ICWS)

Scoring methods differ both in the amount of time that they require of the examiner and in the quality of the information that they provide about a student's writing skills. Advantages and limitations of each scoring system are presented below.

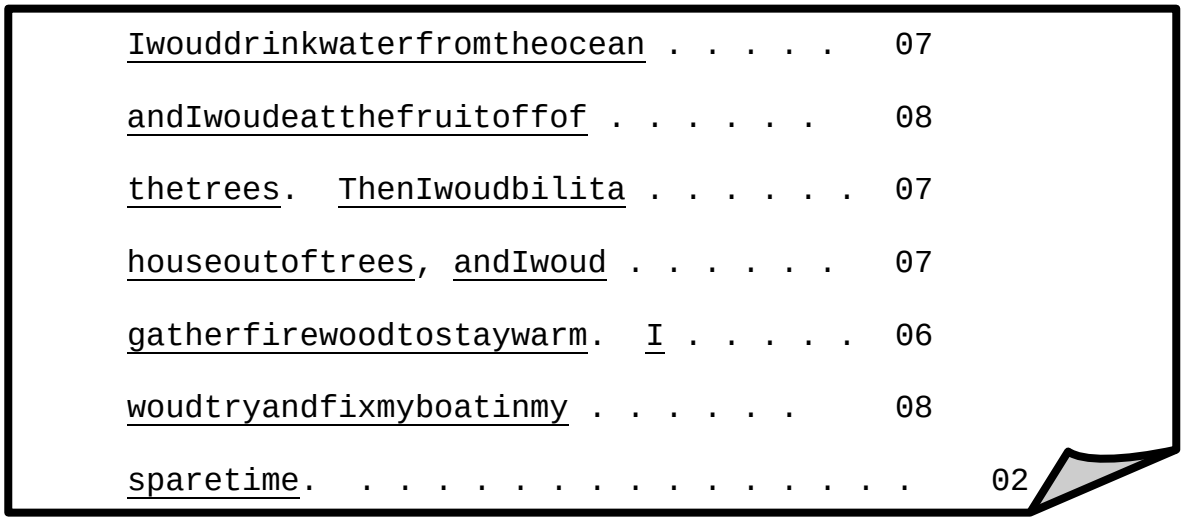
1 – Total Words Written (TWW):

The examiner counts and records the total number of words written during the 3-minute writing probe. Calculating total words written is the quickest of scoring methods. A drawback, however, is that it yields only a rough estimate of writing fluency – that is, how quickly the student can put words on paper – without examining the accuracy of spelling, punctuation, and other writing conventions.

Rules:

a) Any grouping of letters separated by a space is counted.
b) Misspelled words **are counted** in the tally.
c) Numbers written in numeral form (e.g., 5, 17) **are not counted**.
d) The words "The End" **are not counted**.
e) If the student rewrites the story starter, these words **are counted**.
f) **UNDERLINE** each total word written when scoring.

A CBM-WE sample scored for total words written is provided below:



Iwouddrinkwaterfromtheocean 07

andIwoudeatthefruitoffof 08

thetrees. ThenIwoudbilita 07

houseoutoftrees, andIwoud 07

gatherfirewoodtostaywarm. I 06

woudtryandfixmyboatinmy 08

sparetime. 02

Using the total words scoring formula, this sample is found to contain **45 words** (including misspellings).

2 – Correctly Spelled Words (CSW):

The examiner counts up and records only those words in the writing sample that are spelled correctly. Words are considered separately, not within the context of a sentence. Assessing the number of correctly spelled words has the advantage of being quick. Also, by examining the accuracy of the student's spelling, this approach monitors to some degree a student's mastery of written language.

Rules/Considerations:

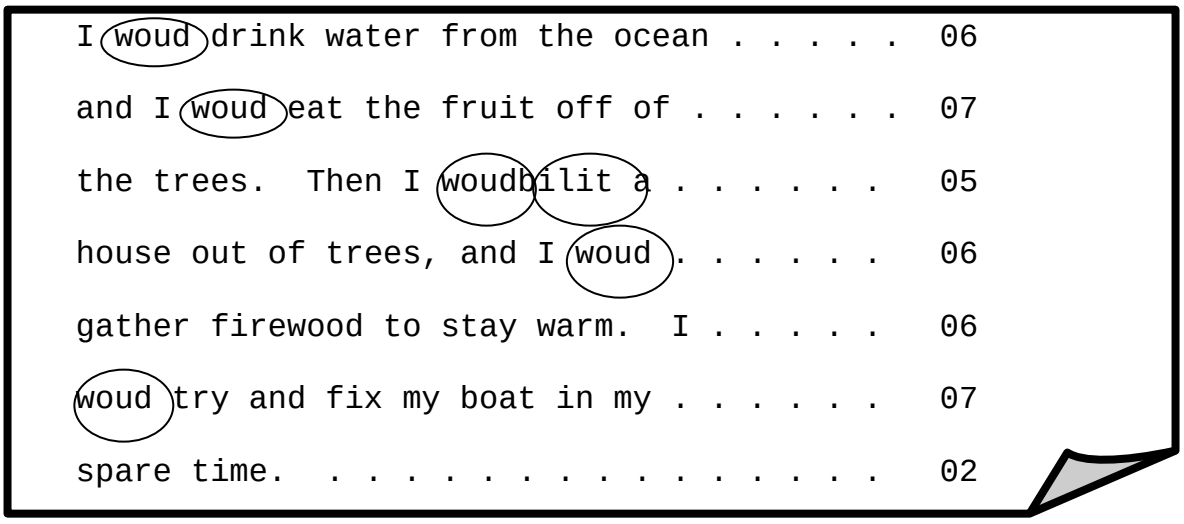
a) When scoring a word according to this approach, a good guideline is to determine whether, in isolation, the word represents a ***correctly spelled term*** in English. If it does, the word **is included** in the tally.

b) For contractions, proper use of apostrophes is ignored. For example, in the sentence, "*That isnt a red car,*" 5 correctly spelled words would be recorded.

c) Assume all names of people **are correctly** spelled.

d) **CIRCLE** incorrectly spelled words.

A CBM-WE sample scored for correctly spelled words is provided below:

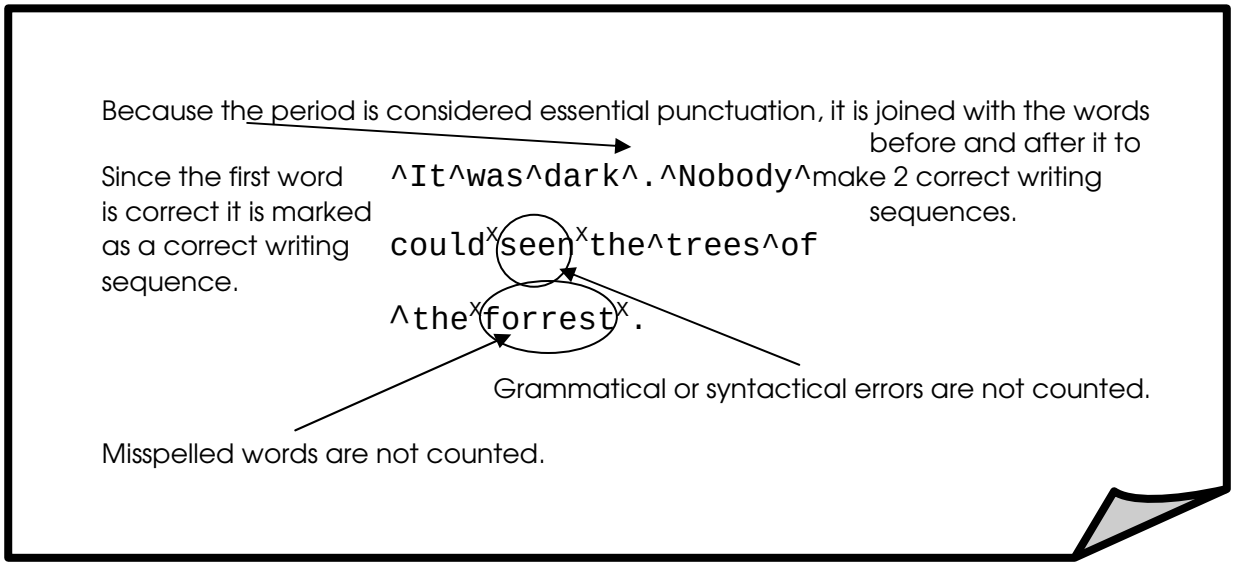


This sample is found to contain **39 correctly spelled words**.

3 – Correct Writing Sequences (CWS):

When scoring correct writing sequences, the examiner goes beyond the confines of the isolated word to consider units of writing and their relation to one another. Using this approach, the examiner starts at the beginning of the writing sample and looks at each successive pair of writing units (writing sequence). Words are considered separate writing units, as are essential marks of punctuation. To receive credit, writing sequences must be correctly spelled, and be grammatically correct. The words in each writing sequence must also make sense within the context of the sentence. In effect, the student's writing is judged according to the standards of informal standard American English. A caret (^) is used to mark the presence of a correct writing sequence.

An illustration of selected scoring rules for correct writing sequences is provided below:



3 – Correct Writing Sequences (CWS):

Rules:

☑ Correctly spelled words make up a correct writing sequence (reversed letters are acceptable, so long as they do not lead to misspellings):

Example: ^Is^that^a^red^car^?

☑ Necessary end marks of punctuation (periods, question marks, and exclamation points are included in correct writing sequences:

Example: ^Is^that^a^red^car^?

All other punctuation, except apostrophes, that is used correctly is counted as well (quotation marks, colons, semicolons, parentheses).

Example: ^Sally^said^,^"^Is^that^a^red^car^?^"

If commas or other punctuation besides the end punctuation is missing, students are **not** penalized for this.

☑ Syntactically correct words make up a correct writing sequence:

Example: ^Is^that^a^red^car^?

^Is^that^a^car^{x}red^{x}?

☑ Semantically correct words make up a correct writing sequence:

Example: ^Is^that^a^red^car^?

^Is^that^a^{x}read^{x}car^?

☑ If correct and capitalized, the initial word of a writing sample is counted as a correct writing sequence:

Example: ^Is^that^a^red^car^?

Capitalization Rule: Only those words that begin a sentence and the word "I" are expected to be capitalized. Do not penalize other capitalization mistakes.

Example:^Is^that^a^Red^ford^car^?

3 – Correct Writing Sequences (CWS):

Rules:

☑ Titles are included in the correct writing sequence count, but not the words "The End":

Example: ^The^Terrible^Rotten^Day

☑ For this measure, numerals **will be counted**.

Example: ^The^14^soldiers^waited^in^the^cold^.

^The^crash^occurred^in^1976^.

Rules:
Not surprisingly, evaluating a writing probe according to correct writing sequences is the most time-consuming of the scoring methods presented here. It is also the metric; however, that yields the most comprehensive information about a student's writing competencies. A CBM-WE sample scored for correct writing sequences is provided below:

^I^X^woud^X^drink^water^from^the^ocean 05

^and^I^X^woud^X^eat^the^fruit^off^of 06

^the^trees^. ^Then^I^X^woud^X^bilit^X^a 05

^house^out^of^trees,^and^I^X^woud 06

^X^gather^firewood^to^stay^warm^.^I 06

woud^X^try^and^fix^my^boat^in^my 06

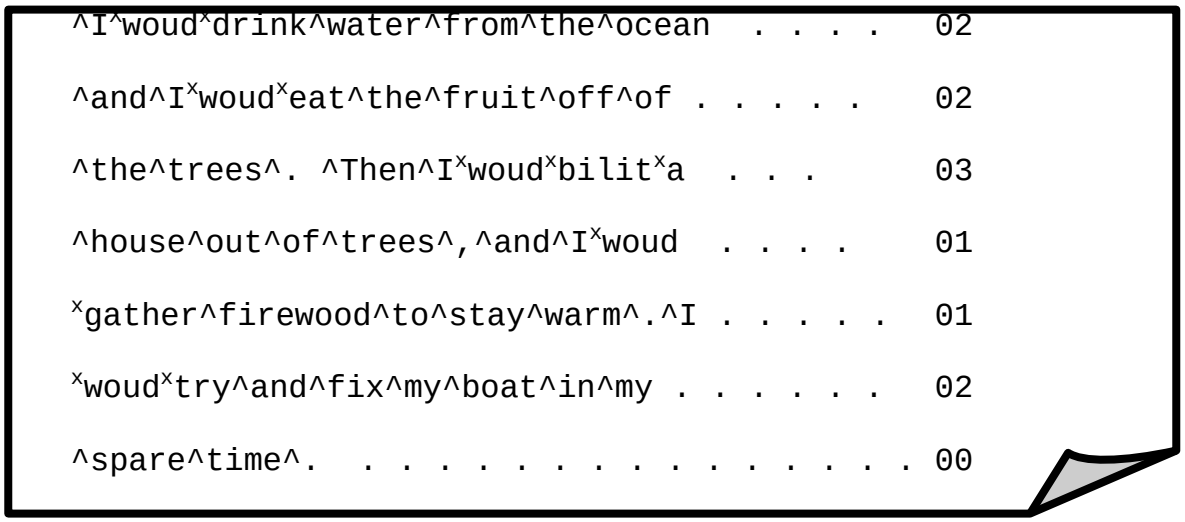
^spare^time^. 03

This sample is found to contain **37 correct writing sequences**.

4 –Incorrect Writing Sequences (ICWS):

This metric further distinguishes writing quality from correct writing sequences. A potential disadvantage of this metric however, is that it not as sensitive to growth in fluency. Counting these sequences can be done simultaneously with correct writing sequences. Any sequence that is not marked by a caret (^) can be marked with an X to designate an incorrect writing sequence. The number of X's can then be tallied.

Here is the same sample with the incorrect writing sequences marked as well:



This sample contains **10 incorrect writing sequences**.

By adding the number of correct writing sequences (i.e., 37) to the number of incorrect writing sequences, we know the total number of writing sequences made was 47.

GENERAL SCORING NOTES

1) Beginning sentences with conjunctions such as `and' & `because' is acceptable.
2) Letter reversals (i.e., writing a letter backwards) should not be penalized.
3) Words that represent sounds (e.g., mmmmm) or create new nouns or names (e.g., a new animal called a catbit) should be counted as correct.
4) If the story ends mid-sentence, this is ok, count correct writing sequences up until the last writing unit but do not count a sequence following the last writing unit.

Example: ^A^red^car

Capitalization

1. ONLY count capitalization as incorrect if capitalization is missing

    a. For the word "I"
    b. Proper names, like Jen or Florida
    c. First word of sentence

2. If you can't distinguish these letters (`c', `w', `m' , `o', `s', 'u', `v', `z') as upper or lower case at the beginning of a sentence, mark it as correct.

3. If a word is capitalized that should not be, just continue like it's correct.

Spelling

1. If a letter is reversed, it is still considered a correctly spelled word (e.g., I bon't like writing).

Hyphens

1. Count a hyphenated word as ONE word (even if it is located in the middle of the sentence).

2. Count the hyphenated word as ONE correctly spelled word (even if it is located in the middle of the sentence).

Punctuation

1. Ignore all incorrect apostrophes.

2. Commas should be given credit when they are used correctly in a series, a date, or to set off punctuation. If used incorrectly, just ignore it.

    a. Example: `^I^like^dogs^,^cats^,^and^X^canaroos^X^.`

    b. Example: `^My^mom,^went^to^school`

Grammar

1. If a word is missing a possessive 's' mark the incorrect sequence but count the word as spelled correctly

    a. Example: `went ^ to ^ grandma ^X^ house`

2. If a verb tense is incorrect, then only count an incorrect sequence for the incorrect noun-verb combination.

    a. Example: `he ^X^ help ^ mom ^ in ^ the ^ kitchen`

Run-On Sentences

1. If the sentence is a run-on sentence, the scorer must decide where the sensible ending is located. Place a vertical line at this point.

    a. Example:^ Murray ^ takes ^ the ^ train ^ to ^ school ^{X}**|**
    ^{X} Mom ^ rides ^ the ^ bus ^ . ^

2. If a run-on sentence is connected by conjunctions, the scorer must determine where to break the sentence apart. Place a vertical line at this point. As a general rule, allow only one or two conjunctions per sentence.

    a. Example: ^ She ^ went ^ to ^ the ^ store ^ and ^ asked ^ for ^ bread ^{X}**|**^{X} and ^{X} looked ^ at ^ books ^ and ^ went ^ home ^ . ^

Spacing Issues

1. If a student separates a word like 'homework' into 'home work', follow the scoring example below:

    a. Example: ^I ^ did^ my^ home ^{X} work^

2. If a student combines 2 words into 1 word, score an ICWS on both sides of the word, for example:

    a. Example: ^There^ were ^{X}alot^{X} of^ pencils

    b. Example: Common mistakes: a lot a few no one

Unfamiliar Names and Slang Words or Phrases

1. Children often make up names in their stories, or use unfamiliar names. In general, do not count a proper name as misspelled unless it's obvious that it is incorrect (e.g., spelling "Sue" incorrectly or misspelling a name that was spelled differently earlier in the passage).

2. Slang words, such as ***gonna***, ***yeah***, ***kinda***, are okay in dialogue only.

3. ***Like*** in the middle of the sentence is incorrect.

   a. Example: `^ He ^ wore ^X like ^X a ^ t-shirt ^ . ^`

Concluding Sentence

1. At the end of the story, the student had to stop writing mid-word. Only count this for total words for the incomplete word.

   a. Example:`^We ^ went ^ to ^ the  sc`

      TWW= 5, CSW = 4, CWS = 4, ICWS = 0

References

Abbott, R. D., & Berninger, V. W. (1993). Structural equation modeling of relationships among developmental skills and writing skills in primary- and intermediate-grade writers. *Journal of Educational Psychology*, *85*(3), 478-508.

Achieve, Inc. (2005). *Rising to the challenge: Are high school graduates prepared for college and work?* Washington, DC: Author.

Agnew, J. L., & Redmon, W. K. (1992). Contingency specifying stimuli: The role of 'rules' in organizational behavior management. *Journal Of Organizational Behavior Management*, *12*(2), 67-76.

AIMSweb®, Inc. (2004). AIMSweb. Retrieved October 20, 2010, from *http://www.aimsweb.com*

AIMSweb® (2010*). AIMSweb growth chart – Written Expression – Letter Sequences Correct and Total Words Written – 2010-2011 (Technical report).* Upper Saddle Ridge, NJ: NCS Pearson.

Ardoin, S. (2006). The Response in Response to Intervention: Evaluating the Utility of Assessing Maintenance of Intervention Effects. *Psychology in the Schools*, *43*(6), 713-725.

Baer, D., Wolf, M., & Risley, T. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, *1*(1), 91-97.

Berninger, V., Abbott, R., Vermeulen, K., Ogier, S., Brooksher, R., Zook, D., et al. (2002). Comparison of faster and slower responders to early intervention in reading: Differentiating features of their language profiles. *Learning Disability Quarterly*, *25*(1), 59-76.

Berninger, V., Cartwright, A., Yates, C., & Swanson, H. (1994). Developmental skills related to writing and reading acquisition in the intermediate grades: Shared and unique functional systems. *Reading and Writing*, *6*(2), 161-196.

Berninger, V. W., Hart, T., Abbott, R. D., & Karovsky, P. (1992).Defining reading and writing disabilities with and without IQ: A flexible, developmental perspective. *Learning Disability Quarterly, 15,* 103-118.

Berninger, V. W., Rutberg, J. E., Abbott, R. D., Garcia, N., Anderson-Youngstrom, M., Brooks, A., & Fulton, C. (2006).Tier 1 and Tier 2 early intervention for handwriting and composing. *Journal of School Psychology, 44,* 3-30.

Berninger, V. W., Vaughan, K. B., Abbott, R. D., Abbott, S. P., Rogan, L. W., Brooks, A., Reed, E., & Graham, S. (1997). Treatment of handwriting problems in beginning writers: Transfer from handwriting to composition. *Journal of Educational Psychology, 89*(4), 652-666.

Berninger, V., Yates, C., Cartwright, A., Rutberg, J., Remy, E., & Abbott, R. (1992). Lower-level developmental skills in beginning writing. *Reading and Writing: An Interdisciplinary Journal, 4,* 257-280.

Betts, J. (2010). Expanding school psychology's horizons for understanding and conducting research: Commentary on the special series on statistical analysis. *Journal of School Psychology, 48*(1), 1-4.

Bramlett, R. K., Murphy, J. J., Johnson, J., & Wallingsford, L. (2002). Contemporary practices in school psychology: A national survey of roles and referral problems. *Psychology in the Schools, 39,* 327-335.

Calfee, R. C., & Miller, R. G. (2007). Best practices in writing assessment. In S. Graham, C. A. MacArthur, & J. Fitzgerald (Eds.), *Best practices in writing instruction* (pp. 265-286) .New York: Guilford.

Campbell, B., Brady, M., & Linehan, S. (1991). Effects of peer-mediated instruction on the acquisition and generalization of written capitalization skills. *Journal of Learning Disabilities*, *24*(1), 6-14.

Campbell, D. & Stanley, J. (1963). *Experimental and quasi-experimental designs for research*. Chicago, IL: Rand-McNally.

Campbell, J., & Willis, J. (1978). Modifying components of creative behavior in the natural environment. *Behavior Modification*, *2*(4), 549-564.

Cooper, J. O., Heron, T. E., & Heward, W. L. (1987). *Applied Behavior Analysis*. Columbus, Ohio: Merrill.

Cutler, L. & Graham, S. (2008). Primary grade writing instruction: A national survey. *Journal of Educational Psychology*, *100*(4), 907-919.

D'Agostino, J. (2005). Measuring learning outcomes: Reliability and validity issues. In Phye, G. D., Robinson, D. H., & Levin, J. R. (Eds), *Empirical methods for evaluating educational interventions* (pp. 112-145). Burlington, MA: Elsevier Academic Press.

Deno, S. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, *52*(3), 219-232.

Deno, S. L., Mirkin, P. K., &Marston, D. (1980).*Relationships among simple measures of written expression and performance on standardized achievement tests*. (Vol. IRLD-RR-22).University of Minnesota, Institute for Research on Learning Disabilities.

Diaz, N., McLaughlin, T., & Williams, R. (1990). The effects of practicing words in sentences on generalization of spelling to written work with mildly mentally handicapped students. *Psychology in the Schools*, *27*(4), 347-353.

Diggle, P. J., Liang, K., &Zeger, S. L. (1994). *Analysis of longitudinal data.* New York: Oxford.

Eckert, T. L., Lovett, B. J., Rosenthal, B. D., Jiao, J., Ricci, L. J., & Truckenmiller, A. J. (2006). Classwide instructional feedback: Improving children's academic skill development. In S. Randall (Ed.), *Learning disabilities: New research* (pp. 271-285). Hauppauge, NY: Nova Science.

Espin, C.A., Skare, S., Shin, J., Deno, S.L., Robinson, S., & Benner, B. (2000). Identifying indicators of written expression proficiency for middle school students. *Journal of Special Education, 34*, 140-153.

Fox, J., & McEvoy, M. (1993). Assessing and enhancing generalization and social validity of social-skills interventions with children and adolescents. *Behavior Modification*, *17*(3), 339-366.

Graham, S. (2006). Strategy instruction and the teaching of writing: A meta-analysis. In C. McArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 187-207). New York: Guilford.

Graham, S., Berninger, V. W., Abbott, R. D., Abbott, S. P., & Whitaker, D. (1997). Role of mechanics in composing of elementary school students: A new methodological approach. *Journal of Educational Psychology*, *89*(1), 170-182.

Graham, S., & Harris, K. (1989). Improving learning disabled students' skills at composing essays: Self-instructional strategy training. *Exceptional Children*, *56*(3), 201-214.

Graham, S., & Harris, K. (1993). Self-regulated strategy development: Helping students with learning problems develop as writers. *The Elementary School Journal, 94*(2), 169-181.

Graham, S., Harris, K. R., Fink-Chorzempa, B., & MacArthur, C. (2003). Primary grade teachers' instructional adaptations for struggling writers: A national survey. *Journal of Educational Psychology*, *95*(2), 279-292.

Graham, S., Harris, K. R., MacArthur, C., & Fink, B. (2002). Primary grade teachers' theoretical orientations concerning writing instruction: Construct validation and a nationwide survey. *Contemporary Educational Psychology, 27*, 147-166.

Graham, S., Harris, K., & 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*(2), 207-241.

Graham, S., & MacArthur, C. (1988). Improving learning disabled students' skills at revising essays produced on a word processor: Self-instructional strategy training. *The Journal of Special Education, 22*(2), 133-152.

Graham, S., MacArthur, C. A., & Fitzgerald, J. (2007). *Best practices in writing instruction.* New York: Guilford.

Guevremont, D., Osnes, P., & Stokes, T. (1988). The functional role of preschoolers' verbalizations in the generalization of self-instructional training. *Journal of Applied Behavior Analysis*, *21*(1), 45-55.

Hammill, D. D. & Hresko, W. D. (1994). *Comprehensive Scales of Student Abilities.* Austin, TX: PRO-ED.

Hammill, D.D. & Larsen, S. C. (1996). *Test of Written Language-Third Edition.* Austin, TX: PRO-ED.

Hanna, E., de Souza, D., de Rose, J., & Fonseca, M. (2004). Effects of delayed constructed-response identity matching on spelling of dictated words. *Journal of Applied Behavior Analysis*, *37*(2), 223-227.

Haring, N. G., Lovitt, T. C, Eaton, M. D., & Hansen, C. L. (1978). *The fourth R: Research in the classroom.* Columbus, OH: Merrill.

Harris, K., & Graham, S. (1996). Constructivism and students with special needs: Issues in the classroom. *Learning Disabilities Research & Practice, 11*(3), 134-137.

Harris, K. R., & Graham, S. (1999). Programmatic intervention research: Illustrations from the evolution of self-regulated strategy development. *Learning Disability Quarterly, 22*, 251-262.

Harris, K. R., Graham, S., Reid, R., McElroy, K., & Stern Hamby, R. (1994). Self-monitoring of attention versus self-monitoring of performance: Replication and cross-task comparison studies. *Learning Disability Quarterly, 17*(2), 121-139.

Harris, K. R., Santangelo, T., & Graham, S. (2008). Self-regulated strategy development in writing: Going beyond NLEs to a more balanced approach. *Instructional Science, 36*, 395-408.

Hattie, J., & Timperley, H. (2007).The power of feedback. *Review of Educational Research*, 77(1), 81-112.

Hayes, J., & Flower, L. (1980). Identifying the organization of writing processes. In L. Gregg & E. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3-30). Hillsdale, NY: Erlbaum.

Hermann, J., Semb, S., & Hopkins, B. (1976). Effects of formal grammar and direct method training on the number of errors in compositions written by sixth-graders. *Revista Mexicana de Análisis de la Conducta*, *2*(1), 68-84.

Hubbard, D.D. (1996). *Technical adequacy of formative monitoring systems: A comparison of three curriculum-based indices of written expression.* Unpublished manuscript, University of Oregon, Eugene.

Kelley, M., & Stokes, T. (1984).Student–teacher contracting with goal setting for maintenance. *Behavior Modification*, *8*(2), 223-244.

Kratochwill, T., & Stoiber, K. (2002). Evidence-based interventions in school psychology: Conceptual foundations of the Procedural and Coding Manual of Division 16 and the Society for the Study of School Psychology Task Force. *School Psychology Quarterly, 17*(4), 341-389.

Kushki, A., Schwellnus, H., Ilyas, F., & Chau, T. (2011). Changes in kinetics and kinematics of handwriting during a prolonged writing task in children with and without dysgraphia. *Research in Developmental Disabilities, 32*(3), 1058-1064.

Locke, E. A., Shaw, K. N., Saari, L. M., & Latham, G. P. (1981). Goal setting and task performance: 1969-1980. *Psychological Bulletin, 90*(1), 125-152.

Lovitt, T. (1975). Applied behavior analysis and learning disabilities: I. Characteristics of ABA, general recommendations, and methodological limitations. *Journal of Learning Disabilities, 8*(7), 432-443.

Macmillan McGraw-Hill (2006). *Treasures grade 3.* New York, NY: Macmillan/McGraw-Hill.

Martens, B. K., Witt, J. C., Elliott, S. N., & Darveaux, D. X. (1985). Teacher judgments concerning the acceptability of school-based interventions. *Professional Psychology : Research and Practice, 16(2),* 191-198.

McGuffin, M., Martz, S., & Heron, T. (1997). The effects of self-correction versus traditional spelling on the spelling performance and maintenance of third grade students. *Journal of Behavioral Education, 7*(4), 463-476.

McMaster, K. L., & Campbell, H. (2006). *Advances in monitoring progress in writing.* Paper presented at the Progress Monitoring Conference, Minneapolis, MN.

McMaster, K. L., & Campbell, H. (2008). New and existing curriculum-based writing measures: Technical features within and across grades. *School Psychology Review, 37*(4), 550-566.

McMaster, K. L., & Espin, C. A. (2007). Technical features of curriculum-based measurement in writing: A literature review. *The Journal of Special Education, 41*(2), 68-84.

Medcalf, J., Glynn, T., & Moore, D. (2004). Peer Tutoring in Writing: A school systems approach. *Educational Psychology in Practice, 20*(2), 157-178.

Minnesota Department of Children, Families, and Learning & NCS Pearson (2002).*Minnesota Basic Skills Test (BST) technical manual for the academic year 2001-2002.*

Mirkin, P. K., Deno, S. L., Fuchs, S. L., Wesson, C., Tidal, G., Marston, D., et al. (1981).Procedures to develop and monitor progress on IEP goals. Minneapolis, University of Minnesota, Institute for Research on Learning Disabilities.

Moher, D., Schulz, K. F., & Altman, D. (2001). The CONSORT statement: Revised recommendations for improving the quality of reports of parallel-group randomized trials. *Journal of the American Medical Association, 285,* 1987-1991.

Monroe, M., & Sherman, E. E. (1966). *Monroe-Sherman Group Diagnostic Reading Aptitude and Achievement Test*. Bradenton, FL: C.H. Nevins.

Monroe, B., & Troia, G. (2006). Teaching Writing Strategies to Middle School Students with Disabilities. *Journal of Educational Research, 100*(1), 21-33.

Moore, W., Heward, W., & Alber, S. (1998). When to self-correct?: A comparison of two procedures on spelling performance. *Journal of Behavioral Education, 8*(3), 321-335.

Moxley, R. A., Lutz, P. A., Ahlborn, P., Boley, N., & Armstrong, L. (1995). Self-recorded word counts of freewriting in grades 1-4. *Education and Treatment of Children, 18*(2), 138-157.

Nihalani, P. K., Mayrath, M., & Robinson, D. H. (2011). When feedback harms and collaboration helps in computer simulation environments: An expertise reversal effect. *Journal of Educational Psychology*, *103*(4), 776-785.

National Commission on Writing. (2003). *The neglected "R."* College Entrance Examination Board.

O'Mahony, P., Dempsey, M., & Killeen, H. (2008). Handwriting speed: Duration of testing period and relation to socio-economic disadvantage and handedness. *Occupational Therapy International*, *15*(3), 165-177.

Parush, S., Pindak, V., Hahn-Markowitz, J., & Mazor-Karsenty, T. (1998). Does fatigue influence children's handwriting performance? Work, 11, 307–313.

Persky, H. R., Daane, M. C., & Jin, Y. (2003). *The Nation's Report Card Writing 2002* (NCES 2003-529). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Powell-Smith, K. A., & Shinn, M. R. (2004). Administration and scoring of written expression curriculum-based measurement for use in general outcome measurement. Eden Prairie, MN: Edformation.

Pratt-Struthers, J., Struthers, T., & Williams, R. (1983). The effects of the Add-A-Word Spelling Program on spelling accuracy during creative writing. *Education& Treatment of Children*, *6*(3), 277-283.

Raudenbush, S.W. & Byrk, A.S. (1988). Methodological advances in analyzing the effects of schools and classrooms on student learning. *Review of Research in Education, 15*, 423-475.

Regan, K., Mastropieri, M., & Scruggs, T. (2005).Promoting expressive writing among students with emotional and behavioral disturbance via dialogue journals. *Behavioral Disorders, 31*(1), 33-50.

Rivera, D., & Bryant, B. R. (1992). Mathematics instruction for students with special needs. *Intervention in School and Clinic, 28*(2*),* 71-86.

Robin, A., Armel, S., & O'Leary, K. (1975). The effects of self-instruction on writing deficiencies. *Behavior Therapy, 6*(2), 178-187.

Sajwaj, T., & Risley, T. (1970). Development and generalization of writing skills in a retarded girl using manipulations of task variables. *Proceedings of the Annual Convention of the American Psychological Association, 5*(Pt. 2), 749-750.

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, Institute of Education Sciences, U.S. Department of Education.

SAS Institute Inc. (2002-2004). SAS 9.1.3 Help and Documentation. Cary, NC: SAS Institute Inc.

Schunk, D., & Swartz, C. (1993). Goals and progress feedback: Effects on self-efficacy and writing achievement. *Contemporary Educational Psychology, 18*(3), 337-354.

Shapiro, E. S. (2004). *Academic skill problems: Direct assessment and intervention*. New York: Guilford.

Singer, J. D. & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence.* New York: Oxford University.

Smith, C. R. (2004). *Learning disabilities: The interaction of students and their environments-Fifth Edition.* Boston: Pearson.

SPSS Inc. (2007). SPSS Base 16.0 for Windows User's Guide. Chicago, IL: SPSS Inc.

Stokes, T., & Baer, D. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10*(2), 349-367.

Stokes, T., Fowler, S., & Baer, D. (1978). Training preschool children to recruit natural communities of reinforcement. *Journal of Applied Behavior Analysis, 11*(2), 285-303.

Stokes, T., & Osnes, P. (1989). An operant pursuit of generalization. *Behavior Therapy, 20*(3), 337-355

Summers, J., & Catarro, F. (2003). Assessment of handwriting speed and factors influencing written output of university students in examinations. *Australian Occupational Therapy Journal, 50*(3), 148-157.

Thorndike, E. L. (1931). *Human learning*. Cambridge: The M.I.T. Press.

Trap, J., Milner-Davis, P., Joseph, S., & Cooper, J. (1978).The effects of feedback and consequences on transitional cursive letter formation. *Journal of Applied Behavior Analysis, 11*(3), 381-393.

Troia, G. (2002). Teaching writing strategies to children with disabilities: Setting generalization as the goal. *Exceptionality, 10*(4), 249-269.

Van Houten, R. (1979). The performance feedback system: Generalization of effects across time. *Child Behavior Therapy, 1*(3), 219-236.

Van Houten, R., Hill, S., & Parsons, M. (1975). An analysis of a performance feedback system: The effects of timing and feedback, public posting, and praise upon academic performance and peer interaction. *Journal of Applied Behavior Analysis, 8,* 449-457.

Van Houten, R., Morrison, E., Jarvis, R., & McDonald, M. (1974). The effects of explicit timing and feedback on compositional response rate in elementary school children. *Journal of Applied Behavior Analysis, 7,* 547-555.

Witt, J. C. & Elliott, J. N. (1983).*Children's Intervention Rating Profile*. Lincoln, NE: University of Nebraska.

Bridget Hier
Vita

Contact:
430 Huntington Hall
Syracuse, NY 13244

Education:

Syracuse University 2009-current
Department of Psychology, School Psychology Ph.D. Program

North Carolina State University 2005-2009
Department of Psychology, B.A.