Syracuse University

SURFACE

Dissertations - ALL
SURFACE

8-2014

Programming Generality into a Performance Feedback Writing Intervention

Bridget Hier
Syracuse University

Follow this and additional works at: https://surface.syr.edu/etd

Part of the Education Commons, and the Social and Behavioral Sciences Commons

Recommended Citation

Hier, Bridget, "Programming Generality into a Performance Feedback Writing Intervention" (2014).
Dissertations - ALL. 153.
https://surface.syr.edu/etd/153

Abstract

Substantial numbers of students in the United States are performing below grade-level expectations in core academic areas, including mathematics, reading, and writing (Aud et al., 2012; National Center for Education Statistics, 2012). National estimates suggest that these deficits are greatest in the area of writing (National Center for Education Statistics, 2012; Persky, Daane, & Jin, 2003), presenting a clear need for research efforts that focus on the development of effective writing interventions. Although performance feedback procedures have been shown to produce promising short-term improvements in elementary-aged students' writing fluency skills (Eckert, Lovett, Rosenthal, Jiao, Ricci, & Truckenmiller, 2006), evidence of maintenance and generalization of these treatment effects is limited (Hier & Eckert, 2014). The purpose of this study was to examine the extent to which programming generality into performance feedback procedures enhanced generality of writing fluency gains. A sample of 118 third-grade students were randomly assigned to one of three conditions: (a) a performance feedback intervention, (b) a performance feedback intervention that incorporated generality programming procedures, or (c) weekly writing practice without performance feedback or generality programming. Intervention effectiveness was assessed in terms of immediate treatment effects, generalization, and maintenance. Results indicated that although the addition of multiple exemplar training to performance feedback procedures did not improve students' writing fluency on measures of stimulus and response generalization, it did result in greater maintenance of intervention effects in comparison to students who received performance feedback without generality programming and students who engaged in weekly writing practice alone.

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

PROGRAMMING GENERALITY INTO A PERFORMANCE FEEDBACK WRITING
INTERVENTION

by

Bridget O. Hier

B.A., North Carolina State University, 2009

M.S., Syracuse University, 2012

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in *School Psychology*.

Syracuse University
August 2014

TABLE OF CONTENTS

Programming Generality into a Performance Feedback Writing Intervention

Within the past decade, the American Psychological Association's Division 16 and the Society for the Study of School Psychology have put forth guidelines that strongly encourage the assessment of generality (i.e., generalization and maintenance) of treatment effects in behavioral research (Kratochwill & Stoiber, 2002). This formal recommendation came after decades of advocacy for such practice by applied behavioral and educational researchers alike. In fact, this issue has been of such interest to applied behavioral researchers that specific, empirically-based techniques for programming and assessing generality of behavioral change were presented in the 1970s and 1980s (Stokes & Baer, 1977; Stokes & Osnes, 1989). Despite calls by prominent researchers and professional organizations to (a) incorporate generality programming techniques into intervention procedures to increase the likelihood of optimal outcomes and (b) evaluate and report the extent to which treatment effects generalize and maintain, intervention researchers have tended to neglect this advice. This lack of attention to generality of behavioral change is particularly evident in the academic intervention literature.

The issue of generality of treatment effects should be of particular interest to those researching academic skills interventions, as recent national data suggest that students in the United States are largely underperforming. Specifically, as of 2011, large percentages of students were unable to demonstrate grade-level proficiency on measures of core academic skills (i.e., mathematics, reading, and writing; Aud et al., 2012; National Center for Education Statistics, 2012). These academic skills deficits were most pronounced in the area of writing, suggesting that there is a current need to develop effective writing interventions that produce lasting and generalizable skill gains. As such, the purpose of this study is to examine the effects

1

of a performance feedback intervention that incorporates generality programming procedures on elementary-aged students' writing fluency (i.e., the ability to write with speed and accuracy). In addition to immediate treatment effects, the effectiveness of this intervention will be evaluated in terms of the extent to which it results in generality of writing outcomes.

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. In their seminal work, Baer, Wolf, and Risley (1968) emphasized that, because generality does not inevitably occur when a behavior changes, applied studies of behavioral change should explicitly examine generality. 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 settings, or (c) extend to other, related, non-trained responses (Baer et al., 1968; Stokes & Baer, 1977). In contrast, generality of behavioral change cannot be claimed when behaviors similar to intervention effects are produced only under conditions similar to those of the intervention itself (Stokes & Baer, 1977). That "generality" and "generalization" are often used synonymously (Cooper, Heron, & Heward, 1987) may mislead one to conclude that any desirable behavior change that occurs in a non-training setting is the result of a single process. As such, it is important to distinguish between the terms. Generality of behavior change can be conceptualized as an overarching process that consists of multiple components (i.e., generalization and maintenance), which are discussed in detail below.

Although the term "generality" was favored by behavioral researchers in the 1970s and 1980s, more recent research has tended to better specify its terminology by distinguishing between generalization and maintenance. In the current review of the literature and the subsequent study, the term "generality" is used to refer to both generalization and maintenance.

2

When referencing one of the components of generality in isolation, the terms "generalization" and "maintenance" are used.

Generalization

Generalization refers to the extension of behavioral change to non-training conditions, and it consists of two behavioral processes: (a) stimulus generalization and (b) response generalization. According to Cooper et al. (1987), stimulus generalization refers to generality across settings, people, and conditions in that it occurs when a learner's performance of a target behavior improves in environments that differ from the original training environment. In general, stimulus generalization is more likely to occur the more a given stimulus resembles the training stimuli. Conversely, when given stimuli configurations differ significantly from stimuli that were present under training or intervention conditions, stimulus generalization is less likely to occur.

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, in a study by Campbell, Brady, and Linehan (1991), student participants demonstrated evidence of response generalization by transferring the skills of identifying words that needed to be capitalized on a worksheet to correctly producing capitalized words in their own writing compositions.

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). This process pertains to the generality of behavioral change across time. The importance of assessing the extent to which a treatment has a lasting effect on behavior (i.e., maintenance) has been thoroughly recognized in the scientific literature for quite some time. In addition to Baer et al.'s (1968) assertion that an essential feature of applied research is the evaluation of a behavior's generality (e.g., the durability of a behavior over time), Lovitt (1975) criticized research that failed to examine the retention of behavioral change and suggested that researchers routinely report retention data.

Importance of Assessing Generality of Treatment Effects

Baer et al. (1968) asserted that although generality is clearly a desired outcome of behavioral change interventions, it is not automatically accomplished when behavioral change is acquired. As such, they argued that assessing for generality is imperative. The recognition of the importance of assessing for generality of treatment effects is not limited to Baer et al.'s seminal advocacy for such regularity. 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). The task force concluded that all treatment studies should provide evidence of the extent to which participants are able to generalize their newly acquired skills across settings and across persons. This, in combination with acquisition data, could then be utilized to determine the effectiveness of a treatment.

Along with generalization, the task force on evidence-based interventions emphasized the importance of assessing for maintenance of intervention effects (Kratochwill & Stoiber, 2002). Specifically, the task force created guidelines for determining the effectiveness of a study's

assessment of maintenance. They concluded 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 results 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 an attempt to encourage intervention researchers within the field of school psychology to evaluate for generalization and maintenance and report their findings, the task force on evidence-based interventions (Kratochwill & Stoiber, 2002) created guidelines outlining the importance and expectations of such practices. However, because behavioral change does not necessitate its generality, researchers may be reluctant to report unsuccessful generality results. For this reason, as well as for optimal participant and client outcomes, it is important to understand methods of programming generality into treatment procedures.

Programming and Assessing Generality of Behavioral Change

Based on their comprehensive review of the literature, Stokes and Baer (1977) described techniques for programming and assessing generalization of behavior change. These researchers found that literature fell into nine categories of techniques used to assess or program generalization: (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 generalization, and (i) reinforce instances of generalization. Although the researchers targeted generalization programming in their literature search, they discovered that in many instances, the aforementioned techniques improved maintenance of intervention effects as well. A majority of literature was categorized under the "Train and Hope" category, suggesting that most studies examining generalization of behavior change did not specifically attempt to program generalization into the treatment; rather, researchers simply trained a behavior to acquisition, then assessed for generality of behavior change. Thus, one limitation of the behavioral literature that was uncovered by this investigation was that much of the reviewed research assessed for generality in the aftermath of an intervention without systematically programming for its occurrence.

In their attempt to refine the work by Stokes and Baer (1977), Stokes and Osnes (1989) reviewed literature that demonstrated promising techniques of programming generalization and maintenance of behavior change to maximize the likelihood of its occurrence. Three overarching categories of generalization programming tactics were found to exist within the literature: (a) exploit current functional contingencies, (b) train diversely, and (c) incorporate functional mediators. For each principle, four tactics were suggested for use in the programming of generalization and maintenance, each of which is described in detail below.

Exploit Current Functional Contingencies

The principle of exploiting current functional contingencies arose from research that successfully promoted generalization and maintenance of behavior change through the

manipulation of a behavior's antecedents and consequences (Stokes & Osnes, 1989). The first

tactic proposed under this category was to train behaviors that are likely to contact powerful

reinforcers in the natural environment. By incorporating reinforcers that are naturally occurring

into the intervention program, it was suggested that newly acquired behaviors would continue to

contact reinforcement upon removal of the training conditions. In the event that reinforcers are

not prevalent enough in the natural environment to promote generalization and maintenance 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 cue their teachers to notice and reinforce their behavior by

asking questions such as, "How is this?" Third, Stokes and Osnes recommended attempting to

decrease the frequency of maladaptive behavior by extinguishing its reinforcing consequences so

that more appropriate behaviors can be developed and maintained. Finally, Stokes and Osnes

argued for the importance of reinforcing instances of generalization as a programming technique.

Campbell and Willis (1978) demonstrated that by using social and token reinforcement for

occurrences of generalization (i.e., variability from trained behaviors), fifth-grade students' essay

writing behaviors increased and maintained over time.

Train Diversely

Stokes and Osnes (1989) argued that because "focused training frequently has focused

effects" (p. 344), using less stringently controlled methods during training conditions likely

promotes generalization and maintenance of behavior change. Training diversely involves

allowing variations in antecedent stimuli, responses, and consequences during training

conditions. One tactic to make training conditions less focused is 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 generalization of

behavior change. Training multiple exemplars does not have to become a resource-heavy task;

even a small number of stimulus exemplars (i.e., two trainers or two settings) are frequently

sufficient to aid in successful generalized outcomes (Stokes & Baer, 1977). Stokes and Osnes

also recommended incorporating into training procedures a subset of responses from the

particular class that is to be generalized and maintained into training procedures (i.e., use

sufficient response exemplars). Third, making antecedents less discriminable may increase the

likelihood that the participant does not perform the desired behavior only under a particular set

of circumstances. Stokes and Osnes also 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.

Incorporate Functional Mediators

Lastly, Stokes and Osnes (1989) suggested that incorporating some type of common

stimulus – possibly a discriminative stimulus – in both the training setting and the generalization

setting may increase the likelihood of successful generalization. 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

generalization and maintenance of behavior change (e.g., Charlop, Schreibman, & Thibodeau,

1985; Marholin & Steinman, 1977; Stokes, Doud, Rowbury, & Baer, 1978). 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 generalization settings has been shown to improve generalized outcomes (Guevremont, Osnes, & Stokes, 1988; Kelley & Stokes, 1984).

Stokes and Osnes's (1989) recommended techniques for programming generalization and maintenance were derived from studies that addressed an extensive array of behavioral outcomes. As such, these techniques have been broadly applied to different types of behavioral interventions, including those that address academic skills. Although behavioral researchers have focused less on developing interventions for academic skills deficits than for other behavioral deficits (Skinner & Daly, 2010), the development of academic interventions that produce generalized effects is an area of great interest among educational researchers.

The Importance of Evaluating Generality of Academic Skills

As previously indicated, in the late 1960s, the maintenance and generalization of behavioral treatment effects were becoming a principal issue among some of the nation's leading psychologists of the time. Following Baer et al.'s (1968) advocacy for the consistent evaluation of generality in behavioral change research, Bandura (1969) claimed that behavioral generalization and maintenance was "by far the most important but most neglected aspect of behavioral change processes" (p. 619). In response to these calls for the assessment of generality of treatment effects, behavioral researchers took strides to increase the reporting of these results in the scientific literature over the next several decades. In fact, applied behavior analysts have made common practice of examining generality of treatment effects; however, this practice continues to be largely neglected by academic skills researchers.

The Education Field's "Central Problem"

The growing interest in ensuring that treatment programs produce lasting, generalized effects on human behavior has been echoed by educational researchers, who have often

9

expressed concern regarding the dearth of academic generality research. For example, Ward and Gow (1982) identified the lack of focus on generalization and maintenance as a "central problem" (p. 231) in the field of educational psychology, and they pointed to an extensive body of research depicting a poor long-term prognosis for individuals with reading difficulties despite a plethora of extant remediation programs demonstrating short-term positive effects. Influenced by seminal works in the area of generality research (e.g., Kazdin & Bootzin, 1972; Stokes & Baer, 1977), Ward and Gow asserted that generalization and maintenance of behavior change should not simply be a coincidental, hoped-for phenomenon within educational research and practice; rather, they argued, generalization and maintenance of behavior change should be explicitly programmed into educational intervention and remediation curricula and should be used as an outcome measure to evaluate a program's effectiveness.

The education field's "central problem" of neglecting to evaluate generality of treatment effects was again exposed in the late 1980s in an empirical examination of then-current issues in the learning disabilities research literature. Bursuck and Epstein (1987) surveyed a nationally representative sample of learning disabilities experts ($n = 66$) to identify their perceptions of the most critical issues that would best inform a future agenda for learning disabilities research and policy. Of the top three issues identified by these experts, the maintenance of treatment effects was one, and the generalization of intervention effects was another. To examine the extent to which published learning disabilities research mapped onto the topics identified as most crucial by the experts, all articles published by two leading learning disabilities journals from the previous year were reviewed and coded according to the issues they addressed. Despite the experts' opinions that generalization and maintenance of intervention effects were among the top three most important issues in the field, only 0% and 0.7% of the research literature was devoted

10

to those topics, respectively. The authors emphasized the importance of assessing for

maintenance and generalization of academic skills gains among the learning disabled population

given teachers' lack of skills or unwillingness to adapt the general classroom environment (i.e.,

generalization setting) to mimic the special education environment (i.e., training setting).

Despite these educational researchers' advocacy for the inclusion of generality evaluation

in the scientific literature, this issue continues to be identified as a problem in the field nearly

three decades later. In an acknowledgment that interventions targeting academic skill deficits

have received far less attention than those targeting behavioral excesses in the applied behavioral

literature, the *Journal of Behavioral Education* published a special issue devoted to research

examining the generalization of academic skills in 2010. In their opening commentary of the

special issue, Skinner and Daly (2010) argued, much like educational researchers from decades

earlier, that research is lacking in the area of generalization of academic skills. That this

continues to be a neglected area of research is particularly alarming in light of recent national

data suggesting that U.S. students' levels of academic proficiency are, on the whole, substandard.

The Current State of Students' Core Academic Performance in the United States

National estimates of students' academic skills in the United States suggest a substantial

need for improvement. The National Assessment of Educational Progress (NAEP), published by

the U.S. Department of Education, reports on national academic data from fourth-, eighth-, and

twelfth-grade students. This organization categorizes student achievement into three levels: (a)

the Basic level denotes partial mastery of skills that are considered prerequisite to grade-level

material; (b) the Proficient level signifies competency with grade-level material; and (c) the

Advanced level represents above-grade-level performance.

Mathematics. Results from the NAEP studies reveal that students' mathematics performance has been steadily increasing since 1996. Specifically, while 21% of fourth-grade students were performing at the Proficient or Advanced levels (i.e., levels that display mastery of grade-level expectations) in 1996, this proportion rose to 36% in 2005 and 41% in 2011 (Aud et al., 2012). A similar pattern was evidenced by the nation's eighth-grade students, with a rise from 24% of students performing at the Proficient or Advanced levels in 1996 to 34% reaching grade-level proficiency in 2011. Although students' mathematics achievement is on the rise, 59% of fourth- and 66% of eighth-grade students continue to perform below grade-level in this area (Aud et al., 2012).

Reading. Unlike the steadily increasing trend in the mathematics data, the NAEP results show a stable trend in students' reading performance over the past decade. For example, while 29% of fourth-grade students demonstrated grade-level proficiency in reading in 1998, this proportion grew to only 34% by 2011 (Aud et al., 2012). Moreover, a substantial percentage of eighth-grade students were unable to read at the Proficient level in 1998, and the same proportion of students (i.e., 67%) continued to perform below grade-level standards on reading measures in 2011.

Writing. Although students' levels of mathematics and reading achievement are clearly suboptimal, results of the NAEP studies suggest that writing is the nation's weakest area in terms of core academic performance. Specifically, 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). By 2011, 73%

of eighth- and twelfth-grade students were unable to demonstrate grade-level proficiency (National Center for Education Statistics, 2012).

Among the fourth-grade students assessed in 2002, 86% of Black students, 83% of Hispanic students, and 86% of American Indian/Alaska Native students performed below the Proficient level, suggesting that writing difficulties were exacerbated for students of minority backgrounds. Although White students and Asian/Pacific Islander students demonstrated higher levels of proficiency, a substantial percentage (i.e., 67% and 59%, respectively) wrote below 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 with grade-level proficiency.

A number of factors may contribute to the writing difficulties experienced by students in the United States. For example, Troia (2002) suggested that students experience inadequate writing instruction in the classroom. In line with this hypothesis, survey data suggest that a substantial percentage of teachers (i.e., 89%) tend not to alter their instructional methods to align with theoretically-relevant practices for struggling writers (Graham, Harris, Fink-Chorzempa, & MacArthur, 2003). Such findings highlight the importance of considering theoretical frameworks that have been proposed as a means of conceptualizing the writing process.

Theoretical Conceptualizations of Writing

Current models of writing are based on cognitive and neuropsychological aspects of the writing process. Indeed, many cognitive and neuropsychological factors have been implicated by researchers as important for the development of written expression. For example, Hooper, Knuth, Yerby, and Anderson (2009) provided a review of the research literature in this area and identified four main cognitive functions that play a role in written expression: (a) visual-spatial

skills, (b) graphomotor output (i.e., planning and control of finger function, orthographic coding), (c) memory, and (d) executive functioning (e.g., organization, planning, attention, initiating, sustaining). Two popular models of writing that emphasize some of these functions are reviewed below.

Hayes and Flower's (1980) Model of Writing

Arguably the most cited model of writing, Hayes and Flower (1980) proposed that the writing process is comprised of three components: (a) planning, (b) translating, and (c) reviewing/revising the written product (see Figure 1). Hayes and Flower proposed that individuals first engage in a period of planning in which they focus on idea generation, organizing, and goal-setting regarding their written work. Following the planning component, they proposed that individuals convert language representations into orthographic text (i.e., translation). Translation was hypothesized to rely heavily on the transfer of ideas in working memory. Finally, upon completing the translation component of the process, Hayes and Flower asserted that individuals review, evaluate, and revise their written products. Importantly, in this model, translation is the only component that is necessary for the act of writing.

Berninger's Model of Writing

Berninger et al. (1997) later argued that the Hayes and Flower (1980) model was more applicable to the adult writing process. They asserted that for children developing emergent writing skills, translation plays a significant role in the writing process. Because planning and revision require higher-level cognitive processes, children who have not yet mastered the lower-level process of text production may have difficulty engaging in these components. Thus, a model was proposed that emphasized the importance of translation for emerging writers such that the component of translation was divided into two sub-components: (a) text generation and

(b) transcription (see Figure 2; Abbott & Berninger, 1993; Berninger, Yates et al., 1992). Text generation is accomplished by 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.

Berninger and colleagues (2006) argued that transcription is directly related to the mechanics of writing (e.g., spelling, punctuation, and grammar), handwriting, and compositional fluency. Two transcription skills in particular (i.e., spelling and handwriting) have been shown to be related to the length and quality of compositions. Specifically, Graham, Berninger, Abbott, Abbott, and Whitaker (1997) administered measures of spelling and handwriting to 600 first-through sixth-grade students to examine the impact of these mechanical variables on compositional fluency and quality. To measure the outcome variables (i.e., writing fluency and quality), participants were also asked to write two compositions. Writing fluency was calculated by counting the total number of words written per 5 minutes, and writing quality was determined by averaging the Likert scale ratings ("1 = considerably below grade expectations; 5 = considerably above grade expectations"; Graham et al., 1997, p. 174) of two experienced teachers who were told to rate the compositions based on content and organization. Results of this study revealed that handwriting and spelling accounted for a significant proportion of the variance observed in writing fluency (41% to 66%) and writing quality (25% to 42%) at both the primary (i.e., first- through third-grade) and intermediate (i.e., fourth- through sixth-grade) grade levels.

Empirical research by Berninger, Cartwright, Yates, Swanson, and Abbott (1994) provided further support that these two sub-components play a significant role in the writing process for children. In this study, 300 fourth- through sixth-grade students completed a large

battery of psychoeducational, neuropsychological, and cognitive tests that included measures of writing skills. Intercorrelations among writing skills were examined with respect to handwriting, spelling, writing fluency, and writing quality. Results indicated that both writing fluency and quality were statistically significantly correlated with handwriting and dictated spelling. Although writing fluency was not significantly related to spontaneous spelling (i.e., spelling within the context of the students' compositions), writing quality was related to this skill. These results suggest that for elementary-aged children, transcription skills (i.e., handwriting and spelling) play a large role in writing ability.

Writing Fluency

The model of writing proposed and supported by Berninger and colleagues' research targets writing fluency (i.e., automaticity and proficiency in transcription) as a fundamental skill in the writing process that should be developed in the elementary grades (Abbott & Berninger, 1993). Developing fluency in the execution of complex skills has typically been conceptualized in the context of automaticity theory, which states that individuals demonstrate automatic, or fluent, responding when they are able to complete a task with speed and accuracy (Samuels, 2006). The study of automaticity has been applied to a variety of domains in the fields of cognitive psychology (e.g., Buckley & Cameron, 2011; Helie, Roeder, & Ashby, 2010; Vachon & Jolicoeur, 2012) and psychoeducation (e.g., Cummings, Dewey, Latimer, & Good, 2011; Frye & Gosky, 2012; Jones & Christensen, 1999). Because fluency is comprised of two components (i.e., speed and accuracy), the measurement of fluent responding involves examining the accurate rate with which a response is produced (i.e., correct responses per an allotted time period; Mathson, Allington, & Solic, 2006). By examining behavioral rate, one is provided with information regarding the speed with which an individual responds.

Like other studies of behavioral fluency, the assessment of writing fluency involves examining the speed and accuracy with which an individual writes within an allotted time frame (Shapiro, 2004). From a cognitive perspective, the ability to write fluently allows children to expend fewer cognitive resources on basic writing components, and thus, more cognitive resources may be utilized for other, higher-level writing components such as composition planning and content knowledge (Graham et al., 1997). Alternatively, from a behavioral perspective, fluent responding is likely to result in improved: (a) retention of a skill, (b) endurance of a skill over longer intervals, (c) stability of a skill in spite of distractions, (d) application of a skill to novel conditions, and (e) spontaneous modification (i.e., adduction) of a skill (see Johnson & Layng, 1996; Martens & Collier, 2011). Thus, children who write fluently are more likely to: (a) retain their skills in the absence of practice, (b) sustain their performance over longer time intervals, even in the face of distractions, (c) transfer their writing skills to novel writing tasks, and (d) modify their writing skills to meet environmental demands.

Authors of the National Writing Project (2003) further implicated the early development of writing fluency as a necessary skill for establishing writing competence. 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). Peverly and colleagues (2007) confirmed that the basic skills of transcription fluency continue to be important in college. Specifically, these researchers found that out of a variety of variables (i.e., verbal working memory, identification of main ideas, and spelling skill), transcription fluency was the only

significant predictor of the quality of lecture notes, which in turn was the only significant predictor of test performance among a sample of 85 undergraduate college students.

Despite the research findings that suggest writing fluency (a) is an important indicator of academic success, (b) should be established in the elementary grades for optimal outcomes, and (c) continues to be a vital foundational skill in the post-secondary setting, substantial evidence suggests that elementary-aged students do not spend a sufficient amount of time practicing writing in the classroom (Cutler & Graham, 2008; Graham et al., 2003). 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 sufficient for the majority of students given that in 2002, 72% of fourth-grade students were unable to write at the Proficient level (Persky et al., 2003) and 43% of referrals for special education services were for students experiencing writing difficulties (Bramlett, Murphy, Johnson, & Wallingsford, 2002). These findings highlight the need for empirically-based interventions that improve students' writing fluency and that result in generality of treatment effects.

Empirically-Based Writing Fluency Interventions

A number of research studies have examined the effects of writing interventions among elementary-aged populations. Over the past several decades, six different types of intervention procedures have been researched with respect to writing fluency (i.e., writing rate and accuracy), proxies of writing fluency (e.g., composition length), or aspects of writing fluency (e.g., spelling accuracy, grammar accuracy, punctuation accuracy). These interventions include: (a) strategy instruction (see Graham, 2006); (b) self-instructional training (i.e., Blandford & Lloyd, 1987; Graham, 1983; Graham & Harris, 1989; Graham & MacArthur, 1988); (c) sentence-combining

18

instruction (i.e., Saddler & Graham, 2005); (d) self-correction (e.g., Cates, Dunne, Erkfritz, Kivisto, Lee, & Wierzbicki, 2007; Jaspers, Williams, Skinner, Cihak, McCallum, & Ciancio, 2012; Moser, Fishley, Konrad, & Hessler, 2012); (e) peer tutoring (i.e., Campbell, Brady, & Linehan, 1991; Maheady, Harper, & Mallette, 1991; Medcalf, Glynn, & Moore, 2004; Watt & Topping, 1993); and (f) performance feedback (i.e., Eckert, Lovett, Rosenthal, Jiao, Ricci, & Truckenmiller, 2006; Hier & Eckert, 2014; Jackson, 1995; Van Houten, 1979; Van Houten, Hill, & Parsons, 1975; Van Houten, Morrison, Jarvis, & McDonald, 1974).

Although there has been a considerable number of writing intervention studies, it is difficult to make definitive conclusions regarding the effectiveness of these interventions. Overall, there are three primary limitations associated with these research studies. First, the vast majority of outcome measures used in this type of research tend to be proxies of writing fluency (i.e., composition length without controlling for task time) or isolated aspects of writing fluency (e.g., accuracy in spelling, punctuation, grammar, or capitalization) rather than metrics that examine both the speed (i.e., writing rate) and accuracy criteria of writing fluency. Second, researchers of writing fluency interventions have not made common practice of systematically programming for and comprehensively assessing generality of treatment effects. Third, much of the literature suggests that even within intervention type (e.g., self-instructional training), these strategies tend to produce variable effects. Exceptions to this third limitation are self-correction and performance feedback interventions, which have been consistently shown to improve writing outcomes. However, with respect to writing, studies examining self-correction interventions have exclusively used this technique in attempts to improve spelling accuracy, which is just one aspect of writing fluency. Therefore, although a great deal of evidence supports the use of self-correction interventions with a broad variety of students (e.g., Cates et al., 2007; Gryiec, Grandy,

& McLaughlin, 2004; Jaspers et al., 2012), thorough analyses of writing fluency interventions should evaluate their effectiveness in terms of writing rate and other measures of accuracy (e.g., spelling in connected text, grammar, punctuation). Additionally, although maintenance of intervention effects has been examined in several of these studies (e.g., Hubbert, Weber, McLaughlin, 2000; Moser et al., 2012; Nies & Belfiore, 2006), evidence of generalization is more limited.

Unlike self-correction interventions, which have only been examined in relation to spelling accuracy, performance feedback interventions have been shown to consistently improve writing rate and accuracy. Additionally, although still limited, some research has been conducted with regard to generalization and maintenance of intervention effects in this domain. Extant research examining the effects of performance feedback writing fluency interventions is reviewed below.

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, feedback exists along a continuum with a clear distinction between feedback and instruction at one end and a complete overlap between the two at the other end. In this sense, when feedback is provided in a correctional manner, it becomes a type of instruction. Feedback that takes on an instructional purpose enables individuals to utilize cognitive and affective processes to reduce the gap between what they currently understand and what is aimed to be understood.

Theoretically, performance feedback has been considered to be cognitive-behavioral in nature. In a study in which adults were explicitly informed whether their guesses to word-

20

meaning associations were correct or incorrect, Thorndike (1931) demonstrated the effectiveness of performance feedback by showing that those who received feedback were more likely to make correct associations compared to participants assigned to a control condition. As such, it has been argued that feedback has the ability to manipulate one's cognitions as well as behavior. For example, Locke, Shaw, Saari, and Latham (1981) 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. Further, Kulhavy (1977) argued against a purely behavioral perspective on performance feedback. He contended that the behaviorist notion that feedback functions as a reinforcer is incorrect. Specifically, he contended that feedback in isolation is not necessarily a reinforcer because the information provided via feedback can be accepted, modified, or rejected by its receiver. Thus, feedback in isolation does not necessitate further action.

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). Indeed, 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.

A great deal is known regarding procedural aspects of performance feedback. For example, Hattie (1999) reported on 74 meta-analyses of performance feedback studies that were conducted with school-aged samples. 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, 0.94 to 1.10). Feedback that included praise, extrinsic rewards, and punishment typically were associated with lower effect sizes (range, 0.14 to 0.31). Additionally, it has been argued that feedback is likely to be most effective when it addresses faulty understanding rather than a complete lack of understanding (Hattie & Timperley, 2007; Kulhavy, 1977). As such, providing a student feedback about a task that is completely unfamiliar to him or her is not likely to impact performance.

Another procedural factor that has been examined with respect to feedback is its timing. Hattie and Timperley (2007) conducted a comprehensive review of the research literature that has examined the impact of feedback on student learning and achievement. Overall, they concluded that most researchers have found evidence supporting the use of more immediate feedback as opposed to delayed feedback. The use of immediate feedback appears to be particularly important as student's work items become more difficult (Clariana, Wagner, & Roher Murphy, 2000).

In addition to the procedural aspects of performance feedback, results of various studies have indicated that the level at which feedback is directed impacts its effectiveness. Hattie and Timperley's (2007) synthesis of the feedback research identified four major levels of focus that feedback may address: (a) feedback about the task, (b) feedback about the processing of the task, (c) feedback about self-regulation, and (d) feedback about the self as a person. Feedback about the processing of the task implies that the feedback provided to the student contains information about learning processes required to accurately complete the task. For example, a teacher may state to a student, "You need to edit this paper by focusing on the grammar." Feedback that is directed toward self-regulation typically involves information that aims to improve self-

evaluation skills or confidence in task completion. For instance, a teacher may say, "You already know the components of a persuasive essay. Now check to make sure you have incorporated all of them into your composition." This information about how to better continue a task with less effort can positively impact self-efficacy and self-regulation. Feedback about the task itself provides information as to whether a student's work is correct (i.e., corrective feedback), and it may direct the student to acquire more or different information (e.g., "Include more information about Benjamin Franklin."). Although feedback about the task is typically less effective for enhancing mastery than feedback about the processes that underlie the task, it is still powerful to the extent to which it contains information that may be used to improve strategy processing or self-regulation. This is consistent with behavioral theory, which suggests that this type of feedback is likely effective because behavior is brought under stimulus control of task instructions (Codding & Poncy, 2010). Additionally, effect sizes for corrective feedback in isolation have been substantial (range, 0.74 to 1.13; Lysakowski & Walberg, 1982; Tenenbaum & Goldring, 1989), further supporting its use. Finally, feedback that is directed toward the self (e.g., "'You are a great student'…'That's an intelligent response, well done'", Hattie & Timperley, 2007, p. 90) tends to be ineffective due to its lack of pertinent information relating to the task, task processes, or self-regulation.

With regard to the four different levels at which feedback can be directed, the extent to which the information is positive or negative can influence student outcomes. In a meta-analysis of feedback interventions, Kluger and DeNisi (1996) argued that both negative and positive feedback can have desired effects on learning behaviors; however, Hattie and Timperley (2007) concluded that the effectiveness differs by feedback level. At the individual level, negative feedback appears to have more potent effects than positive feedback. Negative feedback only

23

tends to be effective at the task level when it is combined with corrective information so that the student acquires knowledge about how to perform under a similar set of conditions the next time. Hattie and Timperley reported that several empirically-based conclusions can be made with regard to feedback directed at the self-regulation level. Specifically, when feedback pertains to the performance of a task that individuals want to do, positive feedback tends to be more effective than negative feedback. Conversely, negative feedback has been shown to be more effective when it relates to a task that individuals have to complete. As such, it is not surprising that positive feedback tends to result in higher rates of persistence toward and interest in a task (Deci, Koestner, & Ryan, 1999).

Although Hattie and Timperley's (2007) comprehensive synthesis provided a great deal of insight into specific aspects of performance feedback, they failed to address several pertinent issues. Specifically, the extent to which aspects of feedback are differentially effective as a function of age, grade level, or skill level is unknown. Further, while the authors discussed feedback in terms of its effects on academic achievement in the general sense, they neglected to discuss its differential effects on outcomes for each content area (e.g., mathematics, reading, writing). Similarly, while Kluger and DeNisi's (1996) meta-analysis showed that feedback had a small, positive effect on performance on average ($d = 0.41$), their results came from a conglomerate of studies with samples consisting of individuals with a variety of ages and skill levels. Additionally, they examined the overall effects of feedback on a large variety of types of performance (i.e., memory retention, reading errors, test performance, performance on puzzles, motor performance, reaction time, arithmetic computations, maintenance jobs, and adherence to regulations). Thus, although Hattie and Timperley (2007) and Kluger and DeNisi (1996)

provided a general overview of performance feedback, their works did not address the specific

effects of performance feedback as a writing intervention.

As it applies to writing, performance feedback has been most frequently researched as a

mechanism to improve writing quality upon revision of a composition. However, several

researchers have used performance feedback to target the less complex skills of writing fluency.

In a series of single-case experimental design studies, Van Houten and colleagues demonstrated

the effectiveness of performance feedback on the outcomes of elementary-aged students

acquiring basic writing skills. Van Houten, Morrison, Jarvis, and McDonald (1974) examined

performance feedback as a component of an intervention package that included explicit timing

and public posting of scores. For this intervention, participants (i.e., 21 second- and 34 fifth-

grade 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. Results of the reversal design showed that the packaged intervention

substantially increased writing rate (i.e., words written per minute) for both second- and fifth-

grade students. 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.

In a follow-up study, 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. Although the element of performance feedback alone doubled the writing

rate of students in comparison to their baseline levels, the combined effects of performance

feedback and public posting of scores increased writing rates by an additional 2.2 words per minute. The effects of teacher praise were variable in that this element 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. Thus, each component (i.e., self-scored performance feedback, public posting of scores, and teacher praise) contributed to the effectiveness of the writing intervention package, but 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) later partially addressed the limitations of his previous studies by examining students' abilities to generalize their writing fluency gains following participation in a performance feedback intervention. Sixty second- through fourth-grade students participated in intervention procedures that involved a combination of self-scored total words written (i.e., performance feedback) and publically posted high scores. Upon completion of the intervention, students' writing rate increased by approximately four words per minute. Stimulus generalization was examined by removing salient intervention stimuli (i.e., students were not told they would receive feedback). Under these conditions, 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). Despite these encouraging findings, performance feedback was not examined in isolation in this study.

In another study that examined performance feedback as part of an intervention package, Jackson (1995) demonstrated the effectiveness of combining performance feedback with public posting of scores and reinforcement to improve students' writing fluency. Specifically, 6 students participated in a single-case design study that used a multiple baseline across behaviors design to sequentially introduce the intervention package to each of three writing behaviors (i.e., writing rate, number of different action verbs, and number of different describing words). Of particular interest is the dependent variable that targeted writing fluency (i.e., writing rate), as this was determined by calculating the total number of words written per 1 minute by each student. During the intervention phase targeting writing rate, students were given 3 minutes to write a story. Upon completion of the 3 minutes, the students self-scored and graphed the total number of words written in their composition (i.e., performance feedback) and were awarded 1 minute of free time for every five words written (i.e., reinforcement). During each session, the amount of free time earned by each student was written on a chart in their classroom (i.e., public posting of scores). The implementation of the packaged intervention procedures yielded an increase in mean writing rate from baseline for all students (range, 15% to 48% increase).

Unlike other writing intervention studies that used performance feedback, Jackson (1995) employed a generalization programming component in which functional mediators were incorporated into the generalization setting. Specifically, during the generalization phase, which assessed the extent to which students' writing rate gains transferred to teacher-generated assignments, students were prompted to count the total number of words written in their compositions. This represented a salient procedure that had been used during training conditions. Importantly, 5 of the 6 participants evidenced increases over baseline in their writing rate on generalization tasks.

Although this study (Jackson, 1995) addressed the general lack of systematic generality programming in the scientific literature, it is limited in several ways. First, maintenance of intervention effects was not assessed; therefore, the extent to which the intervention procedures had a lasting effect on writing behavior in the absence of performance feedback is unknown. Additionally, student participants ranged in age from 11 to 13 years. Thus, although this study provides support for the use of this packaged intervention with adolescents, one cannot generalize these results to a younger population. Finally, like other studies examining the effects of performance feedback on writing fluency skills, the packaged nature of this intervention does not allow for the evaluation of the isolated effects of performance feedback.

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 its isolated effect on writing performance. After writing a story in response to a teacher-selected picture prompt, four fifth- and sixth-grade students with learning disabilities self-scored their compositions 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 composition length of elementary-aged students with learning disabilities, one cannot generalize these findings to a typically-developing population. Additionally, the authors failed to control the amount of time for the writing task, so the metric of writing rate could not be obtained. Furthermore, because maintenance and generalization of writing skill was not examined, one cannot infer whether this type of intervention results in durability of effects over time and transfer of effects across situations.

Results of studies conducted by Eckert and colleagues (2006) address many of the

limitations associated with Van Houten and colleagues' (1974, 1975), Jackson's (1995), and

Harris et al.'s (1994) research. Eckert et al. (2006) randomly assigned 50 third-grade students to

either a control condition or a performance feedback condition. Over a period of 8 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 in

which they were provided with individualized feedback regarding their writing fluency

performance.

Results indicated that students who received performance feedback experienced a

statistically significant improvement in writing fluency ($d = 0.65$, CI = +0.25 to +1.06) and

spelling ($d = 0.74$, CI = +0.33 to +1.14) compared to those students who received practice alone.

Eckert et al. (2006) reported their results of a second study that both further supported the use of

these performance feedback methods to improve elementary-aged students' writing fluency and

showed that the procedures are equally effective regardless of the frequency of their

implementation (i.e., once a week or three times a week; Eckert et al., 2006). Although the

research findings reported by Eckert et al. (2006) suggest that performance feedback in isolation

can be effectively used to increase writing fluency outcomes among typically-developing

elementary-aged students, they too failed to address the extent to which this type of intervention

allows students to generalize and maintain its effects.

In an attempt to address the limitations of their studies, Hier and Eckert (2014) examined

whether Eckert and colleagues' (2006) procedures resulted in greater evidence of maintenance

and generalization for students who received performance feedback compared to those who

received practice alone. In this study, 51 third-grade students received weekly performance

feedback regarding their writing fluency (i.e., total number of words written per 3 minutes) for 6 weeks, and 52 third-grade students were assigned to a practice-only condition. Upon completion of the intervention, Hier and Eckert examined the extent to which students demonstrated evidence of maintenance of intervention effects at 2, 4, and 6 weeks post-intervention. Contrary to expectations, maintenance results favored the practice-only condition. Additionally, although students assigned to the performance feedback condition evidenced greater stimulus generalization (i.e., a transfer of writing fluency gains from probes containing both orally- and visually-presented writing prompts to probes containing only orally-presented writing prompts), no differences between the conditions were found to exist on a measure of response generalization (i.e., a compare-and-contrast essay task that closely resembled students' typical classwork). Thus, 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.

Multiple Exemplar Training

Multiple exemplar training is one procedure that has been used to increase the likelihood that intervention effects generalize and maintain over time. This type of procedure is likely effective in promoting generalization of behavior change because it makes training conditions more varied, thus diversifying stimulus control (Meindl, Ivy, Miller, Neef, & Williamson, 2013;

Stokes & Osnes, 1989). When using this procedure, multiple stimulus and/or response

exemplars may be trained depending on the desired effects.

To date, much of the research supporting the effectiveness of multiple exemplar training

as a means of programming generality has occurred within the context of interventions designed

to increase social behaviors among individuals with autism (e.g., Matson, Sevin, Box, & Francis,

1993; Persicke, Tarbox, Ranick, & St. Clair, 2013; Pollard, Betz, & Higbee, 2012). However,

this technique has also been shown to improve fluent, generalized academic responding. For

instance, Ardoin, Eckert, and Cole (2008) used a multiple exemplar procedure when training

elementary-aged students to improve their oral reading fluency skills. Rather than using a

standard fluency-building procedure in which students read the same passage three times in a

row, students in the multiple exemplar training condition read three different passages that had a

high overlap in word content. Results of that study showed that students who received multiple

exemplar training made greater fluency gains on generalization passages than students who

received the standard repeated readings intervention. Multiple exemplar training has also been

shown to result in greater fluency in matching vocabulary terms to their definitions during

generalization conditions (Meindl et al., 2013).

Multiple exemplar training is likely an optimal procedure that could be used to improve

the generality outcomes of the performance feedback intervention examined by Eckert et al.

(2006) and Hier and Eckert (2014). First, because multiple exemplar training can be effective

without being resource-intensive (Stokes & Baer, 1977), it is a procedure that may be easily

incorporated into school-based intervention research. Second, training multiple response

exemplars directly addresses one of the results found by Hier and Eckert. Specifically, students

who received performance feedback did not evidence better performance on a generalization

measure that mimicked their typical classwork than students who received practice alone.

Although this was the most clinically important task to which the students were expected to

generalize their writing gains, it deviated significantly from training tasks. Thus, by training

multiple response exemplars, training tasks that more readily resemble generalization tasks may

be incorporated into intervention procedures.

Purpose of the Present Study

The purpose of the proposed study was to extend the research literature in the area of

performance feedback writing interventions for elementary-aged students by systematically

incorporating generality programming techniques into Eckert et al.'s (2006) intervention

procedures. Specifically, to increase the likelihood of students maintaining and generalizing

treatment effects of the performance feedback writing intervention, training procedures were

altered to include a programming technique recommended by Stokes and Osnes (1989; i.e.,

training multiple exemplars). Thus, consistent with procedures used by Eckert and colleagues

(2006) and Hier and Eckert (2014), this study included a performance feedback intervention

condition and a practice-only condition. For the purposes of this study, one additional condition

was included to evaluate the effects of incorporating generality programming techniques into the

performance feedback intervention (i.e., performance feedback with generality programming

condition).

There were two aims of the proposed study. The first aim of this study was to evaluate

immediate intervention effects by examining whether students assigned to each of the three

conditions differed in terms of their writing fluency growth trajectories. It was first hypothesized

that all students, regardless of condition, would demonstrate improvements in writing fluency

over time, as both practice and performance feedback have been shown to increase skills in this

area (Eckert et al., 2006; Harris et al., 1994). However, similar to the results reported by Eckert et al. (2006) and Hier and Eckert (2014), it was expected that students who received performance feedback would evidence significantly greater writing fluency gains than students assigned to the practice-only condition. Because the difference in procedures between the performance feedback condition and the performance feedback with generality programming condition targeted maintenance and generalization behaviors rather than immediate treatment effects, it was hypothesized that no difference in growth trajectories would exist between these conditions. Thus, it was hypothesized that students who received performance feedback would demonstrate greater writing fluency improvements than students who received practice alone, regardless of which performance feedback condition they were assigned.

The second, more central aim of the proposed study was to evaluate students' differential performance as a function of condition on (a) a stimulus generalization measure (i.e.., a measure that differed from training in that the presentation of the writing prompt was only orally-presented), (b) a response generalization measure (i.e., a compare-and-contrast essay task that mimicked students' typical classwork), and (c) a maintenance measure (i.e., an assignment that mirrored training tasks but was administered 4 months post-training). Consistent with the results found by Hier and Eckert (2014), it was hypothesized that students assigned to the performance feedback condition would outperform students assigned to the practice-only condition on measures of stimulus generalization, but that no statistically significant difference in performance would exist between students assigned to the performance feedback and practice-only conditions on measures of response generalization. However, given the previous research linking multiple exemplar training with generalized outcomes (e.g., Stokes & Osnes, 1989), it was hypothesized that students assigned to the performance feedback with generality

33

programming condition would demonstrate greater evidence of stimulus and response generalization than students assigned to either of the other two conditions. With respect to maintenance of intervention effects, it was expected that students assigned to the practice-only condition would demonstrate greater evidence of maintenance of treatment effects than students assigned to the performance feedback condition. This hypothesis was based upon the results of Hier and Eckert's (2014) research, as students who received practice alone were more likely to maintain their fluency skills from the final intervention session than students who received performance feedback. The researchers suggested that these results were theoretically explainable because (a) procedures of the maintenance session were exactly the same as intervention sessions for students who received practice alone, whereas the salient feedback stimulus that had been present during intervention was removed for students who were assigned to the performance feedback condition and (b) students assigned to the practice-only condition were likely "maintaining" lower levels of performance. Finally, because multiple exemplar training is also linked with improvement in maintenance (e.g., Stokes & Osnes, 1989), it was hypothesized that students assigned to the performance feedback with generality programming condition would demonstrate greater evidence of maintenance than students assigned to the performance feedback or practice-only conditions.

Method

Participants and Setting

Third-grade students from one public elementary school located in a rural, northeastern school district were recruited to participate in the study after obtaining approval from Syracuse University's Institutional Review Board. Students for whom parental permission (Appendix A) and student assent (Appendix B) were obtained were screened for eligibility and invited to

participate in the study. Students were considered eligible for participation in the study if: (a) they were not experiencing severe motor deficits that precluded them from composing written stories; (b) they were not experiencing severe cognitive deficits that resulted in eligibility for special education services; (c) English was their primary language; (d) they were not classified as having a Learning Disability; (e) they did not have a one-to-one instructional aide or a Section 504 plan specifying additional instructional modification; (f) they demonstrated minimum proficiency by writing at least seven words on a baseline measure (described in Measures); and (g) they legibly scribed a subset of letters from the alphabet. Ineligible students and those students without consent participated in an alternate instructional activity assigned by their classroom teacher.

After excluding students who did not meet the eligibility criteria, a total of 118 third-grade students were included in this study (38 in the control condition, 37 in the performance feedback condition, and 43 in the generalization programming condition) (see Figure 3). Table 1 displays the participants' demographic data. The average age of student participants was 8.5 years ($SD = 0.49$), and no students included in the sample received special education services. Results indicated that the majority of students identified as White (97.5%), whereas few participants identified as American Indian or Alaska Native (0.8%), Asian (0.8%), or Black or African American (0.8%). Additionally, 100% of students in both conditions identified with the same ethnicity (i.e., not Hispanic or Latino). The sample consisted of slightly more males (57.6%) than females (42.4%).

All participating teachers ($N = 7$) held at least a bachelor's degree, and four (57%) held master's degrees in education or developmental sociology. One teacher (14%) held an additional

certification in literacy. The mean number of years of teaching experience was 14.1 (range, 1.5 to 28 years).

The school was 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. According to the most recent New York State School Report Card, which published demographic data for the 2011-2012 school year, the participating school was comprised of 542 third- through sixth-grade students. Of the 542 students enrolled in this school, 38% were eligible for free or reduced-priced lunch. The majority of students were identified as White (96%), and a substantially smaller percentage were identified as Asian or Native Hawaiian/Other Pacific Islander (1%), Black or African American (1%), and Hispanic or Latino (1%).

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) provided 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.

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 response generalization. Secondary measures, specifically, the Essay Composition subtest of the Wechsler Individual Achievement Test – Third Edition (Pearson, 2009), 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. Student and teacher intervention acceptability measures 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-WE, students were provided with one probe containing a story starter (e.g., "I was talking to my friends when all of a sudden ...") and were instructed to spend 1 minute planning a story based on the story starter. Students were then instructed to spend 3 minutes writing a narrative story, and were prompted by the assessor to continue writing for the entire 3 minutes.

A total of 10 probes were used over the course of this study (i.e., one probe for baseline assessment, one probe during each intervention session, two probes for pre- and post-stimulus generalization measures, and one probe during the maintenance session). Each probe contained a story starter that has 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. A complete listing of story starters that were used in this study is provided in Table 2.

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 students previewing 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.

38

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 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 3. 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).

Response generalization probe. Students were administered a writing task that resembled their typical classwork in the subject of writing, as identified by their classroom teachers (Appendix G). The exact nature of this compare-and-contrast writing task was determined upon contact with classroom teachers during pilot work by Hier and Eckert (2014). In that study, all classroom teachers identified the Treasures (Macmillan McGraw-Hill, 2006) compare-and-contrast test as the end-of-year exam that all third-grade students were required to complete, and as such the teachers reported frequently administering past exams as classroom writing assignments. Although none of the teachers in the current study used the Treasures

series, they all reported that the Treasures compare-and-contrast essay task was similar to the writing assignments they typically assigned to their students based on the New York State Common Core curriculum.

To examine the extent to which writing fluency gains transferred to this writing assignment, a modified version of the Treasures compare-and-contrast test was administered. Using a procedural script (Appendix H), the experimenter visually and orally presented the students with an essay topic (see Table 4). 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.

Generalization training probe. Students assigned to the performance feedback with generality programming condition were asked to complete three generalization training probes during the course of the intervention phase (see Appendix G). The generality training probes were abbreviated versions of the response generalization probe. Specifically, this measure differed from the response generalization measure in two ways: (a) rather than having 3 minutes to plan their composition, students had only 1 minute to plan and (b) rather than having 10 minutes to write their composition, students had only 3 minutes to write. These training probes were intended to explicitly program generality into the performance feedback intervention by providing multiple exemplar training. Specifically, by alternating these probes with CBM-WE probes during intervention sessions, multiple exemplars were trained. Table 4 provides a list of the essay topics provided to students for each generalization training probe.

Wechsler Individual Achievement Test – Third Edition. The Wechsler Individual Achievement Test – Third Edition (WIAT-III; Pearson, 2009) is a standardized, norm-referenced

test designed to measure academic abilities of children aged 4 to 19. The WIAT-III is comprised of 16 subtests designed to measure academic competence in the areas of oral language, reading, mathematics, and written expression. For the purposes of this study, the Essay Composition subtest of the WIAT-III was administered to examine students' spontaneous writing abilities. This subtest required students to plan and write an essay based on a verbally-provided writing prompt (i.e., "Write about your favorite game…") for 10 minutes. At the onset of administration of this task, students were encouraged to (a) include at least three reasons for their enjoyment of the game and (b) try to write a full page of text. The essay was then evaluated in the areas of (a) Word Count (i.e., written productivity), (b) Theme Development and Text Organization (e.g., organizational quality, message clarity, idea elaboration), and (c) Grammar and Mechanics (i.e., grammar, capitalization, spelling, punctuation).

The test manual of the WIAT-III reports the technical adequacy of the measure, which has been primarily evaluated by the test developers (Pearson, 2009). In terms of reliability, the Essay Composition subtest has been shown to have high test-retest reliability ($r = .88$) among 8- and 9-year-old children, and correlation coefficients for interscorer agreement were greater than .90 for each of the evaluation criteria (i.e., Word Count, Theme Development and Text Organization, and Grammar and Mechanics). In addition, performance on the Grammar and Mechanics subtest has been shown to reliably differentiate students with and without a Specific Learning Disability in the area of written expression.

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 to provide an initial indicator of orthographic skill (Appendix I).

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 were asked to print a set of 10 lowercase letters from the alphabet (i.e.,

f, c, r, m, v, y, i, h, e, o; Appendix J). 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.

Student intervention acceptability measure. An acceptability measure was

administered to all students following the intervention phase to assess their perceptions of the

procedures used in the study. The questions used in this measure were adapted from the

Children's Intervention Rating Profile (Witt & Elliott, 1983) and have been used in previous

studies examining students' levels of acceptability of academic interventions. For instance, Hier

and Eckert (2014) administered these questions to a sample of 92 third-grade students upon

receiving intervention in writing fluency over a 6-week period, and the measure was found to

have acceptable internal consistency among both the experimental condition (8 items, $\alpha = .77$)

and the control condition (6 items, $\alpha = .71$).

This assessment consisted of a series of questions that each used a 5-point Likert-type

response system. Response options ranged from "*not at all*" to "*very, very much.*" Students

assigned to the performance feedback and performance feedback with generality programming

conditions received a five-page packet containing eight questions regarding their attitudes

towards writing, their perceptions of procedures used in the intervention, and their perceptions toward receiving feedback (Appendix K). The first four and last two questions were also administered to the students assigned to the practice-only condition. Cronbach's alpha reliability coefficients were calculated and the measure was found to have adequate internal consistency for the practice-only condition (6 items, $\alpha = .76$) and the performance feedback condition (8 items, $\alpha = .79$). The measure demonstrated good internal consistency in the generality programming condition (8 items, $\alpha = .86$).

Intervention Rating Profile-15. 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 L) to measure their acceptability of intervention procedures. The abridged version of this rating scale consisted of 15 questions that were 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. Results of the current study revealed excellent internal consistency for the adapted measure (15 items, $\alpha = .90$).

Teacher questionnaire regarding writing instruction. Participants' classroom teachers were asked to complete the Writing Orientation Scale (Graham, Harris, MacArthur, & Fink, 2002; Appendix M) 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 the writing instruction provided in the participants' classrooms was also measured and is reported below.

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 are significantly correlated with associated writing practices (range, $r = .17$ to $.31$). The internal consistency of the scale (i.e., alpha) is .70, demonstrating an acceptable level of 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 four phases: (1) an eligibility and baseline assessment phase spanning 2 weeks; (2) an intervention phase spanning 6 weeks; (3) a generalization assessment phase spanning 1 week; and (4) a maintenance assessment phase spanning 4 months. 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. Students who were ineligible to participate in the study completed classroom instructional materials developed by the classroom teacher.

A maintenance assessment session was conducted by research assistants 4 months following the last intervention session, and this span of time included students' summer vacation (i.e., 2 months). At that time, 92% ($n = 108$) of the full sample was available for follow-up assessment. Of the 10 students who were not retained for maintenance assessment, no patterns

appeared to exist in terms of their condition assignment (i.e., 40% were assigned to the practice-only condition, 30% were assigned to the performance feedback condition, and 30% were assigned to the generality programming condition), sex (60% were males), or race (100% were White).

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 M). Research assistants collected the questionnaires from the teachers during the eligibility and baseline assessment phase.

Eligibility and baseline assessment phase. At the beginning of the first session, students were administered an assent form (see Appendix B). Upon being read a statement describing the nature of their involvement in the study, students were asked to circle "Yes" if they would like to participate or "No" if they would not like to participate. Next, 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 J). 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

students 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 fewer 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. The Essay Composition subtest of the WIAT-III (Pearson, 2009) was administered in a group format. Students were visually and verbally provided with a writing prompt (i.e., "Write about your favorite game…"), and then were given 10 minutes to plan and write a story in response. Finally, participants were administered the pre-intervention response generalization task.

Following this eligibility and baseline assessment phase, students were randomly assigned to a performance feedback condition, a performance feedback with generality programming condition, or a practice-only condition. Procedures for each of these conditions are described below, and Appendix N displays a schematic summary of the procedures.

Individualized performance feedback condition. Students assigned to the performance feedback and performance feedback with generality programming 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 was

complete, students completed a CBM-WE probe for the remainder of the session.

        Individualized performance feedback with generality programming condition. To

assess the extent to which explicitly programming generality into the performance feedback

intervention resulted in greater generality of treatment effects, modifications were made to the

individualized performance feedback intervention procedures for students assigned to the

performance feedback with generality programming condition. Although students received

performance feedback in an identical fashion to those assigned to the performance feedback

condition (i.e., a box showing total words written and an arrow), the writing probe they received

alternated weekly between CBM-WE probes and generalization training probes. Specifically,

during the first week of the intervention phase, students received a CBM-WE probe after

receiving performance feedback. The following week, students received a generalization

training probe upon receiving performance feedback. The required task (i.e., CBM-WE probe or

generalization training probe) continued to alternate by week throughout the duration of the

intervention phase. During sessions in which students received CBM-WE probes, procedures

were identical to those of the performance feedback condition. However, during weeks in which

students received generalization training probes, procedures differed slightly from those of the

performance feedback condition with respect to the task that was presented following performance feedback.

Using a procedural script (see Appendix O), the research assistant provided students with task instructions. Similar to the performance feedback condition, the first page of the packet for students assigned to the performance feedback with generality programming condition contained students' identifying information. The next page of the student packet contained performance feedback in such a manner that is identical to that of the performance feedback condition. Students were then asked to complete a CBM-WE probe or a generalization training probe for the remainder of the session.

Practice-only condition. Procedures in the practice-only condition were the same as those of the individualized performance feedback condition, but the individualized performance feedback step was omitted. Thus, student packets in this condition contained 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. Following the last intervention session, student participants and their classroom teachers were asked to complete a measure of their perceptions of the intervention (see Appendices K and L, respectively). To ensure students' accurate understanding of the questions, the research assistant guided them through each question and their associated response options 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 the week following the final intervention session. This probe

differed from the training CBM-WE probe in that rather than visually displaying a story starter at the top of the response sheet, research assistants orally presented students with a story starter. With the exception of this probe change 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).

Response generalization session. Students were asked to complete a response generalization task the day before the stimulus generalization task. The purpose of the response generalization task was to examine students' abilities to transfer writing fluency gains to a task that was similar to their typical classroom writing assignments. Similar to the CBM-WE probes, writing prompts were both visually and orally presented for this task. However, 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 session. Four months after the final intervention session, which included students' summer vacation, students were administered one CBM-WE maintenance 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, the maintenance task was identical to intervention sessions.

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 an indication of students' writing quality. Furthermore, correct writing sequences have been shown to be the most accurate measure of

49

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. A detailed scoring manual is provided in Appendix R.

Secondary measures. The number of total words written on each CBM-WE probe was calculated to include on individualized performance feedback forms and to determine students' instructional levels. 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). The metric of total words written provides an indication of students' writing fluency performance in comparison to an established criterion for their grade level. 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. Conversely, 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 WIAT-III (Pearson, 2009) are reported below. 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, performance feedback with generality programming, or practice-only condition.

Procedural Integrity

To assess the extent to which study procedures were conducted with fidelity, a permanent product measure was completed by secondary research assistants for 68.5% ($n = 61$) of 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 credit was given accordingly. Procedural integrity was calculated by dividing the instances when the secondary research assistant indicated the primary research assistant accurately implemented a step by the total number of possible procedural steps and multiplying that by 100%. Results indicated that 100% of steps were accurately completed by the primary research assistants for each observed session. Table 5 lists the procedural integrity data for each condition.

Interscorer Agreement

Following initial data scoring, a random selection of 33% of CBM-WE probes was selected for interscorer agreement. These probes were re-scored for the primary dependent measure (i.e., correct writing sequences), and any instance of disagreement was re-examined by the primary researcher to make a final decision regarding the score. Interscorer agreement was calculated by dividing the number of agreements by the sum of agreements and disagreements. The mean interscorer agreement was 99.1% (range, 85.71% to 100%). To account for errors in

agreement due to chance, Kappa coefficients were also calculated. The mean Kappa coefficient was .98 (range, .65 to 1), indicating very high agreement in the scoring of correct writing sequences.

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 by another trained research assistant 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, calculate generalization and maintenance analyses, and conduct secondary analyses. A hierarchical linear modeling function in SAS was used to examine students' writing fluency progress over time.

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 -1and 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

Student demographics. As previously reported, descriptive statistics were calculated for the entire sample and by condition (see Table 1). Nonparametric analyses were conducted to determine whether significant differences in these demographic variables existed between

conditions. Results revealed no significant differences between conditions with respect to sex, χ^2 (1, $N = 118$) = 1.57, $p = 0.46$, or race, χ^2 (4, $N = 118$) = 6.18, $p = 0.40$. In addition, parametric analyses revealed no differences in mean age between conditions, F (2,115) = 0.27, p = .76. These findings suggest that on average, the students assigned to each condition were similar in terms of their demographic variables.

Students' baseline writing skills. To describe students' baseline writing skills, descriptive statistics were calculated for students' performance on the WIAT-III (Pearson, 2009), the Paragraph Copying Task, and the first CBM-WE probe (see Table 6). Importantly, 11% ($n = $ 13) of student participants wrote fewer than 30 words on the WIAT-III Essay Composition subtest. Because essays containing fewer than 30 words cannot be reliably scored (Pearson, 2009), those students' WIAT-III scores were not included in the analyses. Analyses of variance (ANOVAs) with Scheffe's test specified were computed to determine whether differences existed between conditions in the mean scores of each measure. Results indicated that no statistically significant differences in performance existed between conditions for correct writing sequences on the first CBM-WE probe, F(2, 117) = 1.08, $p = .34$, total words written on the first CBM-WE probe, F(2, 117) = 0.71, $p = .50$, Essay Composition standard scores on the WIAT-III, F(2, 104) = 0.29, $p = .75$, Grammar and Mechanics standard scores on the WIAT-III, F(2, 104) = 1.04, $p = .36$, or the number of words correctly copied on the Paragraph Copying Task, F(2, 117) = 0.23, $p = .80$.

The relationship between students' performance on each of the baseline measures was examined using Pearson correlations (see Table 7). Standard scores on the WIAT-III Essay Composition measure were strongly correlated with standard scores on the Grammar and Mechanics scale ($r = .66$, $p < .01$) and were moderately correlated with the number of words

correctly copied on the Paragraph Copying Task ($r = .42$, $p < .01$) and correct writing sequences ($r = .46$, $p < .01$) and total words written ($r = .33$, $p < .01$) on the baseline CBM-WE probe. Although standard scores on the WIAT-III Grammar and Mechanics scale were also moderately correlated with the number of total words written on the baseline CBM-WE probe ($r = .32$, $p < .01$), they were strongly correlated with the number of words correctly copied on the Paragraph Copying Task ($r = .51$, $p < .01$) and the number of correct writing sequences on the baseline CBM-WE probe ($r = .58$, $p < .01$). The number of correct writing sequences on the baseline CBM-WE was strongly correlated with the number of total words written ($r = .87$, $p < .01$) and the number of words correctly copied on the Paragraph Copying Task ($r = .55$, $p < .01$). Scores on the Paragraph Copying Task were moderately correlated with the total number of words written on the baseline CBM-WE probe ($r = .48$, $p < .01$).

Teachers' writing orientations and instructional practices. Students' experiences with writing in the classroom were examined by conducting descriptive analyses for the results of the questionnaire that was administered to teachers at baseline. Table 8 summarizes the teachers' collective responses on the Writing Orientation Scale (Graham et al., 2002). Results of this measure indicate that the teachers' writing philosophies most strongly aligned with explicit instruction (M factor score = 5.0; $SD = 0.86$), and they placed less emphasis on natural learning (M factor score = 4.05; $SD = 1.22$) and correct writing (M factor score = 3.40; $SD = 1.26$). Six (86%) of the teachers reported using the 6+1 Trait® writing model (Culham, 2005) to guide their classroom writing instruction. Of those six teachers, three reported using additional writing curricula and techniques in the classroom (i.e., New York State Common Core Writing Standards, Lucy Calkins Program [Calkins, 1994], Four Square Writing Method [Gould, 1999], Houghton Mifflin© writing program [Houghton Mifflin, 2013]). With respect to the frequency

of teachers' specific instructional practices, few teachers reported providing daily instruction in component writing skills (e.g., spelling, handwriting, planning and revising); however, more teachers reported providing this instruction on a weekly basis (see Table 9). When asked to estimate the amount of time students spend practicing writing in the classroom, the teachers reported that students spend a weekly average of 140.7 minutes ($SD = 208.5$ minutes; range, 15 minutes to 600 minutes) practicing spelling, 122.1 minutes ($SD = 211.3$ minutes; range, 30 minutes to 600 minutes) composing written work, and 38.2 minutes ($SD = 30.4$ minutes; range, 10 minutes to 100 minutes) practicing handwriting.

Major Analyses

Analyses were conducted to assess (a) whether performance feedback significantly improved students' writing fluency over time (i.e., immediate treatment effects), (b) whether students were able to maintain gains in writing fluency over a period of time, and (c) whether the students evidenced generalization of their writing skills to different writing formats.

Immediate treatment effects of performance feedback. Using the metric of correct writing sequences, the trajectory of students' writing fluency growth (i.e., slopes) throughout the duration of the intervention phase was examined for the performance feedback condition, the performance feedback with generality programming condition, and the practice-only condition. Multilevel modeling for repeated measures (PROC MIXED function in SAS V9.3 software, SAS Institute, 2014) was used to analyze the between-condition differences in students' slopes. A Level 1 analysis was specified to estimate the patterns of intra-individual growth by a linear model, and a Level 2 analysis was specified to examine the between-condition differences in the intercept (i.e., estimated baseline performance) and slope (i.e., rate of change in performance across sessions). Because multilevel models use a reference group, a single model could not

make the three comparisons needed for the purposes of this study. Specifically, by computing

one model, two of the three conditions would not be directly compared. This would require a

minimum of two models to be computed with different reference groups in each, leading to one

comparison being made twice. Thus, to avoid redundancy, and for the purpose of model

parsimony, three two-group conditional growth models were computed to examine pairwise

comparisons.

First, an empty model containing only the intercept was examined to determine the

intraclass correlation coefficient (ICC), which is a measure of between-person variability. The

ICC was calculated using the intercept and residual estimates that were produced by the empty

model. Results indicated that approximately 54.93% of the total variance in this model was

explained by between-person variability, which supports the use of multilevel modeling as an

appropriate method of analysis for these data (Lee, 2000).

To produce the Level 1 model (i.e., unconditional growth model), intervention session

was added to the empty model. The time variable (i.e., session) accounted for a substantial

amount of variance (*pseudo R^2* = .18). Results from this model suggest that participants

demonstrated significant gains in correct writing sequences across intervention sessions, with an

average gain of 1.48 correct writing sequences per session, t (117) = 8.86, p <.001.

To examine whether statistically significant differences existed in students' writing

fluency growth by condition, three conditional growth models were produced that each contained

the additional variable of condition. The conditional growth models allowed for the following

pairwise comparisons in writing fluency growth: (1) the practice-only condition and the

performance feedback condition, (2) the practice-only condition and the performance feedback

with generality programming condition, and (3) the performance feedback condition and the

performance feedback with generality programming condition.

The variables of sex and baseline instructional level were first examined to determine

whether they should be entered as predictor variables into the conditional growth models. Sex

did not explain a significant amount of variance in any of the models, and was therefore not

included in the final conditional growth models: (1) t (350) = -1.59, p = .11; (2) t (381) = -1.79, p

= .07; (3) t (373) = -1.52, p = .13. Baseline instructional level did explain a significant amount

of variance in all three conditional growth models: (1) t (350) = 3.21, p = .001; (2) t (381) = 5.31,

p < .001; (3) t (373) = 6.06, p < .001. Therefore, baseline instructional level was included as a

predictor variable in the final conditional growth models.

The first conditional growth model examined the difference in slopes between the

practice-only condition and the performance feedback condition (see Table 10). Results

indicated that although students from each condition performed similarly at baseline, t (350) =

0.30, p = .77, students assigned to the performance feedback condition evidenced statistically

significantly greater writing fluency growth throughout the course of the intervention than

students assigned to the practice-only condition, t (350) = 2.36, p = .02. Specifically, students

assigned to the practice-only condition gained an average of .90 correct writing sequences per

week whereas students who received performance feedback gained an average of 1.89 correct

writing sequences per week. Figure 4 displays the predicted values from this model.

A second conditional growth model compared the difference in slopes between the

practice-only condition and the generality programming condition (see Table 10). Parameter

estimates revealed no statistically significant differences in correct writing sequences between

the conditions at baseline, t (381) = -0.81, p = .42. Students assigned to the generality

programming condition made greater gains than those assigned to the practice-only condition with an average increase of 1.64 correct writing sequences per session. However, this difference was only marginally significant, t (381) = 1.92, p = .056. Figure 5 displays the predicted values from this model.

The final conditional growth model was computed to examine the difference in slopes between the performance feedback condition and the generality programming condition (see Table 10). Results of this model indicated that students performed similarly at baseline, t (373) = -1.15, p = .25. Additionally, no statistically significant difference in writing fluency growth over time was found to exist between the performance feedback condition and the generality programming condition, t (373) = -0.60, p = .55. Figure 6 displays the predicted values from this model.

Generalization of treatment effects. To examine whether students' writing fluency differed significantly from one another on measures of generalization as a function of condition (i.e., performance feedback, performance feedback with generality programming, and practice-only), two one-tailed analyses of covariance (i.e., ANCOVAs) were computed with alpha set at .05.

The first ANCOVA was computed for the stimulus generalization measure to analyze whether significant differences existed in the mean number of correct writing sequences on the post-intervention measure as a function of group assignment while controlling for performance on the baseline stimulus generalization measure. Upon examining the assumptions of ANCOVA, it was determined that was an appropriate analysis to use in this case. Specifically, the covariate (i.e., correct writing sequences on the baseline stimulus generalization probe) and the dependent variable (i.e., correct writing sequences on the post-intervention stimulus

generalization probe) were significantly and strongly correlated with one another, $r = .70$, $p <$

.001. Visual inspection of the scatterplot (Tabachnick & Fidell, 2007) indicated that the

covariate and dependent variable were linearly related. In addition, results from a univariate

analysis of variance revealed that the assumption of homogeneity of regression slopes was

retained, F (1, 107) = .44, $p = .65$. Finally, a Levene's test of equality of error variances revealed

that the assumption of homogeneity of variance was upheld, $F = .29$, $p = .78$.

Table 11 displays the results of the ANCOVA, which suggest that the differences in

correct writing sequences between conditions on the post-intervention stimulus generalization

probe were statistically significant when controlling for baseline writing fluency. Bonferroni

post hoc tests were conducted on all possible pairwise contrasts. Results of the post hoc analyses

revealed that students assigned to the performance feedback with generality programming

condition ($M = 32.94$, $SD = 14.43$) wrote significantly more correct writing sequences on the

post-intervention stimulus generalization measure than students assigned to the practice-only

condition ($M = 27.53$, $SD = 13.77$), $p = .03$. No statistically significant difference in mean

correct writing sequences existed between the performance feedback condition ($M = 31.34$, $SD =$

13.03) and the practice-only condition, $p = .16$. Contrary to the initial hypothesis, the mean

difference in correct writing sequences between the performance feedback and performance

feedback with generality programming conditions was not statistically significant, $p = .50$.

Further, the effect of condition on correct writing sequences was small for the stimulus

generalization measure, partial $\eta^2 = .05$. Overall, these results suggest that students who were

exposed to explicit generality programming during training were no more likely to generalize

their writing fluency gains to a task that differed from the training task in terms of writing

prompt presentation than students who received the same intervention without generality

programming. Notably, although students assigned to both performance feedback conditions demonstrated modest improvements in their writing fluency on the stimulus generalization measure from pre- to post-intervention, students in the practice-only condition demonstrated a decline.

A second ANCOVA was conducted to determine whether statistically significant differences in the mean number of correct writing sequences existed between conditions on the post-intervention response generalization measure while controlling for performance on the baseline response generalization measure. Examination of the assumptions revealed that an ANCOVA was an appropriate analysis to use for the response generalization data.   The covariate (i.e., correct writing sequences on the baseline response generalization probe) was significantly and moderately correlated with the dependent variable (i.e., correct writing sequences on the post-intervention response generalization probe), $r = .59$, $p < .001$. Visual inspection of the scatterplot supported the assumption of linearity between the covariate and dependent variable. Similarly, the assumption of homogeneity of regression slopes was also upheld, $F (1, 109) = .29$, $p = .75$. Finally, a Levene's test supported the assumption of homogeneity of variance, $F = .17$, $p = .84$.

The response generalization results are listed in Table 12. Results did not support the initial hypothesis, as no statistically significant differences in mean writing fluency performance were found to exist between any of the conditions when controlling for performance on the baseline response generalization measure. Additionally, the effect of condition on response generalization was small, partial $\eta^2 = .04$. Interestingly, the trend in the response generalization data was similar to that of the stimulus generalization data in that (a) the mean number of correct writing sequences for the performance feedback and performance feedback with generality

programming conditions were very similar and (b) the mean number of correct writing sequences for the practice-only condition was substantially lower than the other two conditions. These results suggest that participation in an intervention that incorporated an evidence-based generality programming component (i.e., multiple exemplar training) did not result in improved generalization to a task that resembled students' typical classroom writing assignments.

Maintenance of intervention effects. Maintenance of intervention effects was assessed for 92% ($n = 108$) of the total sample. To examine maintenance, the percentage of gains and/or losses in correct writing sequences was calculated. Specifically, for each individual, the percent gain or loss was computed by subtracting the score on the final training probe from the score on the maintenance probe, dividing this number by the score on the final training probe, and multiplying by 100. Student participants were considered to have maintained their treatment effects when they obtained a percent change score of 0 or greater, representing no decline in writing fluency from the final intervention session. Table 13 lists the descriptive statistics, mean percent change scores, and percentage of students who demonstrated maintenance of intervention effects in each condition.

A chi-square analysis was used to examine whether significant differences in the proportion of students who evidenced maintenance of intervention effects existed between conditions. Results indicated that the proportion of students who maintained their performance from the final intervention session was moderately associated with the condition to which they were assigned, χ^2 (2, $N = 108$) = 6.03, $p = .05$, Cramer's V = .24. An examination of the standardized residuals revealed that fewer than expected students in the performance feedback condition evidenced maintenance of intervention effects ($z = -1.3$), but a greater than expected number of students in the generality programming condition maintained intervention effects ($z =$

1.0). However, neither of these standardized residual values surpassed the critical value of $z =$ 1.65, which would have represented statistical significance in a one-tailed test. The standardized residual for the practice-only condition ($z = .2$) indicated that only slightly more students than expected maintained their performance from the final intervention session.

Overall, results of the maintenance analyses suggest that explicitly programming for generality during intervention resulted in greater maintenance of treatment effects over a 4-month period. Specifically, students assigned to the performance feedback condition experienced a decline in their writing fluency on the maintenance probe (mean percent change = -7.4%). However, for students who received generality programming in addition to performance feedback, an improvement in writing fluency was observed from the final intervention session to the maintenance session (mean percent change = 18.5%). In addition, generality programming resulted in a higher percentage of students who maintained intervention effects than performance feedback alone and practice alone.

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) students' and teachers' acceptability of intervention procedures.

Shifts in instructional level. To examine the clinical significance of the intervention effects, 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 the final intervention session. Results indicated that shifts in instructional level were significant following intervention for all three conditions, $\chi^2_{MB} = 48.74$, df = 3, $p < .001$. When each condition was analyzed separately, results indicated that both the performance

feedback condition, $\chi_{MB}^2 = 22.0$, df $= 3$, $p < .001$, and the generalization programming condition, $\chi_{MB}^2 = 17.57$, df $= 3$, $p = .001$, experienced significant shifts in instructional level from pre-intervention to the final intervention session. Although the practice-only condition also demonstrated shifts in instructional level (see Table 14), the statistical significance of this movement was unable to be analyzed due to one of the cells (i.e., baseline instructional level) failing to meet the minimum required count of 1.

An additional McNemar-Bowker test was computed to examine shifts in instructional level from the final intervention session to the maintenance session. Although the student participants were in fourth-grade at the time of the maintenance assessment, their third-grade instructional level equivalents were used to allow for comparison to their final intervention performance. No significant shifts in instructional level were observed for any of the conditions (practice-only condition: $\chi_{MB}^2 = 1.13$, df $= 3$, $p = .77$; performance feedback condition: $\chi_{MB}^2 = 5.58$, df $= 3$, $p = .13$; generality programming condition: $\chi_{MB}^2 = 1.17$, df $= 3$, $p = .76$) or the sample as a whole, $\chi_{MB}^2 = 1.13$, df $= 3$, $p = .77$.

Overall, these results suggest that both performance feedback interventions resulted in statistically significant shifts in students' instructional levels immediately following treatment. However, from the final intervention session to the maintenance session, students in all conditions tended to remain in the same instructional range.

Table 14 displays the percentage of students classified at each instructional level at baseline, the final intervention session, and the maintenance session across the entire sample and by condition. Although 87.3% of students performed within the frustrational range of writing fluency at baseline, the majority of students assigned to the performance feedback (64.9%) and the generality programming (53.5%) conditions completed the intervention phase writing in the

instructional or mastery range. Conversely, the majority (57.9%) of the students assigned to the

practice-only condition remained in the frustrational range at the final intervention session. The

mastery range in Table 14 denotes performance in the instructional or mastery level for fourth-

grade students. Thus, when maintenance was assessed, only 36.4% and 44.1% of students in the

practice-only and performance feedback conditions, respectively, demonstrated grade-level

competence in writing fluency. Notably, a substantial percentage of students in the performance

feedback condition evidenced less fluent writing than they had during the final intervention

session. In contrast, students assigned to the generality programming condition demonstrated

improvement from the final intervention session to the maintenance session with 51.2% of

students performing within the third-grade mastery range (i.e., at least instructional level for

fourth-grade) at follow-up. Ultimately, 45.5% of students in the practice-only condition, 47.1%

of students in the performance feedback condition, and 39.0% of students in the generality

programming condition continued to write at the frustrational level.

      Student acceptability of intervention procedures. Students' acceptability of the

intervention procedures are reported descriptively in Table 15. Students' overall acceptability

ratings were moderate for the practice-only condition ($M = 3.65$, $SD = 0.97$), the performance

feedback condition ($M = 4.00$, $SD = 0.89$), and the generality programming condition ($M = 4.17$,

$SD = 0.86$). To examine whether significant differences existed in students' overall acceptability

as a function of the condition to which they were assigned, an ANOVA was computed with

Scheffe's tests specified. Results indicated that significant differences in acceptability ratings

existed between conditions $F(2, 111) = 3.27$, $p = .04$. Post hoc analyses revealed that although

no significant difference existed between the two conditions that received performance feedback

(p = .27), students assigned to the generality programming condition rated the intervention

procedures significantly higher than students who received practice alone (p = .05).

Teacher acceptability of intervention procedures. Teachers' acceptability of the

intervention procedures as measured by the modified version of the IRP-15 is reported

descriptively in Table 16. Overall, the teachers rated the procedures as highly acceptable (M =

5.31, SD = 0.38), with greater acceptability ratings for the following aspects of the intervention:

(a) the appropriateness of the intervention for the students' writing difficulties (M = 5.67, SD =

0.52), (b) their willingness to use the intervention within the classroom context (M = 5.67, SD =

0.52), (c) the likelihood that the intervention would not result in adverse side effects (M = 5.67,

SD = 0.52), (d) the procedures used (M = 5.67, SD = 0.52), and (e) the likelihood that students

would benefit from the intervention (M = 5.67, SD = 0.52). However, slightly lower teacher

ratings were obtained regarding the extent to which their students experienced severe enough

writing difficulties to warrant use of the intervention (M = 4.83, SD = 0.41). Further, there was

more variability in their reports of how consistent the intervention procedures were with those

they have used previously (M = 4.33, SD = 1.63).

Discussion

The purpose of this study was to evaluate the effectiveness of performance feedback and

generality programming procedures as a means of improving elementary-aged students' writing

fluency outcomes. It was hypothesized that, in accordance with previous research (e.g., Eckert et

al., 2006; Hier & Eckert, 2014), providing students with feedback regarding their writing rate

would increase their fluency over time. Indeed, results of the current study supported this

hypothesis, as students assigned to the performance feedback condition demonstrated statistically

significantly greater fluency growth over the course of a 6-week intervention than students who

practiced writing without feedback. A second hypothesis was that incorporating multiple

exemplar training into the performance feedback procedures would improve students' abilities to

maintain their writing fluency gains over time and generalize those gains to different writing

formats. Results of this study revealed that although the combined performance feedback and

multiple exemplar training procedures did not improve students' abilities to generalize their

writing skills to a greater extent than performance feedback alone, they did result in greater

maintenance of intervention effects over a 4-month period. The primary findings of this study

are discussed in further detail below.

Effects of Performance Feedback and Generality Programming on Students' Writing

Fluency Growth

Previous research has demonstrated that both practice and performance feedback can

significantly improve elementary-aged students' writing fluency (Eckert et al., 2006; Harris et

al., 1994; Hier & Eckert, 2014). Because all students in this study received either practice alone

or a combination of practice and performance feedback, it was expected that students in all

conditions would demonstrate improvements in writing fluency over time. This hypothesis was

confirmed, as the total sample of students gained an average of 1.48 correct writing sequences

per session over the course of the study.

It was also predicted that students who received performance feedback would evidence

greater writing fluency growth through the course of the intervention than students who received

practice alone, as was demonstrated in previous studies (Eckert et al., 2006, Hier & Eckert,

2014). Because the generality programming procedures targeted generalization and maintenance

rather than immediate treatment effects, the addition of those procedures was not expected to

improve students' writing fluency to a greater extent than performance feedback alone. The

findings supported this hypothesis, as multilevel modeling results revealed no statistically significant differences in growth slope estimates between the two performance feedback conditions. Specifically, students assigned to the generality programming condition gained an average of 1.64 correct writing sequences per session, which was comparable to the mean gain of 1.89 correct writing sequences per session for students assigned to the performance feedback condition. In comparison to the practice-only condition's average gain of 0.90 correct writing sequences per week, the performance feedback and generality programming conditions evidenced greater fluency growth over time. Importantly, the average growth in each condition was higher than national reporting standards, which suggest that third-grade students who do not receive formal writing intervention gain an average of 0.30 correct writing sequences per week (AIMSweb, 2010).

Students who received weekly performance feedback regarding their writing fluency also demonstrated more positive shifts in their instructional level than students assigned to the practice-only condition. Specifically, although 87% of students began the intervention writing within the frustrational range, the majority of students who received performance feedback (65%) and generality programming (54%) ended the 6-week intervention in the instructional or mastery range. A smaller percentage of students (42%) in the practice-only condition wrote within the instructional or mastery range during the final intervention session. These findings are supported by previous studies (e.g., Hier & Eckert, 2014; Koenig, 2013), which have found performance feedback to result in greater positive instructional shifts than practice alone. Interestingly, Hier and Eckert (2014) reported nearly identical results for a sample of urban students who were exposed to the same performance feedback procedures described in this

study; however, far fewer of the students in the urban sample reached the instructional or mastery level (i.e., 22%) after receiving practice alone.

Consistent with previous research (Hier & Eckert, 2014), results of this study indicated that across each condition, students' baseline instructional level was a significant predictor of their writing fluency growth in response to the intervention. Specifically, students who began the intervention by writing in the frustrational range were more likely to demonstrate greater growth rates than students who wrote in the instructional or mastery ranges. Although these results could suggest that the performance feedback intervention may be less appropriate for students who are already fluent writers, only 15 students (12.7%) were writing in the instructional or mastery range prior to intervention. Thus, these results should be interpreted cautiously.

Effects of Generality Programming on Generalized Outcomes

Because previous research has indicated that performance feedback procedures may produce only short-term improvements in writing fluency with limited evidence of generalization to clinically meaningful writing tasks (Hier & Eckert, 2014), one of the main purposes of this study was to examine the effects of generality programming on students' generalization of writing fluency gains. Following 6 weeks of intervention, students were administered a measure of stimulus generalization and a measure of response generalization. It was hypothesized that the addition of generality programming would improve students' generalized responding on both measures to a greater extent than feedback alone. Overall, results of this study did not support the initial hypothesis, as students who received generality programming in addition to performance feedback did not perform differently than students who received performance feedback alone on either measure of generalization.

On the stimulus generalization measure, when controlling for baseline performance, students assigned to the two performance feedback conditions performed similarly whereas students assigned to the practice-only condition demonstrated less fluent responding. Although there were not statistically significant differences in the performance of students assigned to the performance feedback condition in comparison to those who received practice alone, the trend in the data was similar to previous findings (Hier & Eckert, 2014) in that students in the performance feedback condition gained correct writing sequences from pre- to post-intervention while students in the practice-only condition lost correct writing sequences. Additionally, consistent with the initial hypothesis, there were statistically significant differences in the performance of students who received generality programming in comparison to those students who received practice alone. These findings are consistent with previous research (Silber & Martens, 2010), which found multiple exemplar training in the area of reading to result in significantly greater fluent responding on generalization probes in comparison to a control condition, but not in comparison to a previously established reading fluency intervention that did not explicitly incorporate generality programming.

The trends in students' performance on the response generalization measure were similar to the trends observed on the stimulus generalization measure. Specifically, when controlling for baseline performance, students assigned to the performance feedback and generality programming conditions performed nearly identically on the response generalization measure, while students assigned to the practice-only condition performed substantially lower. Despite the weaker performance of the practice-only condition, no statistically significant differences in performance on the response generalization measure existed between any of the conditions. This finding was consistent with previous research (Hier & Eckert, 2014), which found that students

who received practice alone and feedback alone performed similarly on measures that mimicked typical classwork following intervention. However, results of the current study conflicted with the initial hypothesis that generality programming would result in an improved transfer of intervention effects to a clinically meaningful measure of students' writing.

Although multiple exemplar training has been shown to improve generalized responding across stimulus and response conditions (e.g., Marzullo-Kerth et al., 2011; Matson et al., 1993), there are several potential reasons why that did not occur in the context of this study. First, aspects of the response generalization measure may have hindered students' abilities to demonstrate generalization of writing fluency gains following intervention. Specifically, given the compare-and-contrast nature of the response generalization measure, students were required to utilize background knowledge in their story compositions, which was not controlled. This task requirement could have plausibly affected performance on the measure, as previous research has indicated that there is substantial individual variability in students' background knowledge, particularly in the elementary grades (Hirsch, 2003).

An additional aspect of the response generalization measure that could have affected this study's results was the time requirement associated with the task. Research from the occupational therapy literature suggests that writing for 10 minutes, as was required on the response generalization measure, can result in handwriting fatigue for third-grade students (Parush, Pindak, Hahn-Markowitz, & Mazor-Karsenty, 1998) that significantly decreases handwriting speed over the course of the writing period (Kushki, Schwellnus, Ilyas, & Chau, 2011). Thus, high rates of text production over an extended time period may not have been a physical possibility for the students in this study. In this sense, the extended length of time that

students were required to write on the response generalization measure could have resulted in ceiling effects that masked differences in their writing fluency.

A broader explanation for the lack of impact of the generality programming procedures on students' generalized responding concerns the method of analysis. Because generalization of intervention effects is typically examined in the context of single case experimental designs, the scientific literature consists of very few exemplars of appropriate analytical methods for within-group designs. Thus, without a standard means of analyzing generalization in the literature, it is possible that the method of analysis used in this study was inadequate. One study that did examine generalized responding in the context of a group design used a slightly different method of analysis than was used in the current study. Specifically, Silber and Martens (2010) used an ANOVA to examine the difference in reading fluency gains from pre- to post-intervention between conditions. Thus, for each individual participant, a difference score was computed in the number of words read correctly per minute from pre- to post-intervention. This differed from the current study in that baseline performance was examined relative to post-intervention performance at the individual level rather than the group level. Future research should aim to resolve this problem that is currently present in the group-design research literature (i.e., lack of a standard approach for analyzing generalization of intervention effects).

In addition to aspects of the measurement and method of analysis of generalization, another potential explanation for the observed results involves the manner in which generality was programmed. Specifically, due to the length of the intervention phase, students assigned to the generality programming condition received 50% less practice on CBM-WE training probes than students assigned to the other two conditions. Similarly, they were given only three instances of exposure to the generality training probes, which may not have provided them with

sufficient practice to generalize their skills. Thus, more time spent in the intervention phase may have produced more desirable effects on skill generalization for students who received generality programming.

A second point related to the generality programming procedures is that although multiple exemplars were trained, the exemplars used in the current study may not have been sufficient to produce generalized effects. Specifically, although Stokes and Baer (1977) suggested that two exemplars are typically sufficient in training conditions to lead to generalization, that may not be the case for the complex skill of writing. Additionally, conditions are considered sufficient to facilitate generalization when the stimuli are similar or share some common characteristic. However, the task instructions between the training condition and the generalization condition sampled different stimuli and responses. Thus, the difference in the task instructions between the training conditions and generalization conditions may have been ample enough to hinder successful skill generalization.

A final consideration for the generalization results observed in this study involves the relationship between maintenance and generalization. As noted by Martens and Eckert (2007), it is unclear whether the ability to generalize skill gains across stimuli and responses is dependent upon a prerequisite ability to generalize skills across time, as limited models of generalization have been experimentally evaluated. Some researchers have suggested that the relationship between maintenance and stimulus and response generalization is hierarchical such that maintenance must be established before generalization can occur (Haring, Lovitt, Eaton, & Hanson, 1978; Rivera & Bryant, 1992). Thus, it is possible that the intervention concluded before students' skills were sufficiently developed to a level that would allow for stimulus and response generalization following exposure to diverse practice.

Effects of Generality Programming on Maintenance of Intervention Effects

Because multiple exemplar training has been shown to improve students' maintenance of treatment effects (e.g., Marzullo-Kerth, Reeve, Reeve, & Townsend, 2011), it was hypothesized that students who were exposed to this procedure would show stronger evidence of maintenance over a 4-month period than students who received practice or performance feedback alone. Indeed, 63% of students in the generality programming condition evidenced maintenance of writing fluency gains whereas only 54% and 35% of students in the practice-only and performance feedback conditions maintained their gains. This finding was not likely due to one condition's underperformance during the final intervention session, as no statistically significant differences in the mean number of correct writing sequences existed between conditions. These results suggest that within the generality programming condition, stimulus control over students' fluent responding was sufficiently strengthened to an extent that produced generalization of intervention effects across time.

One technique that has been consistently established as an effective method for programming for maintenance of intervention effects is the inclusion of naturally occurring reinforcers as part of treatment (Foxx, 2013). Importantly, this technique may have been inadvertently incorporated into the generality programming procedures. Specifically, students in the generality programming condition received feedback regarding their writing performance on an abbreviated measure that was similar to their typical classwork. Thus, it is possible that for students in this condition, the feedback mimicked a naturally occurring consequence (e.g., teacher feedback) that was sufficient to maintain higher rates of writing fluency over time.

Although diverse practice (i.e., multiple response exemplar training) serves to diversify stimulus control and thus promote response generalization more so than it targets response

73

endurance (i.e., maintenance), there is evidence to suggest that multiple exemplar training alone, in the absence of procedures that are associated with targeted maintenance programming (e.g., intermittent reinforcement, incorporating naturally occurring consequences, fading contingencies), can improve temporal generalization. For instance, Marzullo-Kerth et al. (2011) used multiple exemplar training in the absence of more targeted maintenance programming procedures in an intervention designed to increase sharing behaviors in children with autism. Although the multiple exemplar training was incorporated into the intervention to increase the likelihood of response generalization, it also resulted in maintenance of intervention effects.

Limitations

Several limitations were present in the study. First, with respect to internal validity, the threat of diffusion could not be controlled in this study. Because students across seven classrooms were randomly assigned to one of the three conditions, it is possible that the unique procedures of each condition were revealed to students in different conditions upon returning to the classroom. Additionally, because all sessions were administered in group format, social competitions in performance between students may have developed over the course of the study. In addition, due to the sample that was used, external validity threats were also present in this study. Specifically, participants represented a homogeneous group of third-grade students from one small, rural school. Thus, the results of this study should not be extrapolated to different populations. Finally, because the procedures of this study were conducted under a high degree of experimental control, one cannot generalize these findings to less controlled contexts (e.g., implementation by teachers as part of typical classroom procedures).

Directions for Future Research

Given the results of this study, there are several avenues for future research. Broadly, this study presents a need for the relationship between maintenance, stimulus generalization, and response generalization to be experimentally evaluated to determine in what contexts maintenance must be established before generalization can occur. Additionally, the impact of adapting the procedures used in this study to more systematically manipulate stimulus control as a function of student responding should be examined. For instance, rather than initially diversifying stimulus control as was done in this study through the use of multiple exemplar training, it may be beneficial to first strengthen stimulus control through fluency training procedures (e.g., performance feedback). Then, once students have established a functional fluency level (i.e., writing within instructional level; Codding, Archer, & Connell, 2010), intervention procedures should shift to focus on generality programming.

Future research should also compare different generality programming techniques to determine which are the most effective in the context of the performance feedback intervention. For instance, as previously discussed, it is unclear whether programming naturally occurring consequences into the generality programming condition's procedures resulted in, enhanced, or had no effect on students' maintenance outcomes. Thus, future researchers should consider examining each of these techniques in isolation. Additionally, examining the effectiveness of these procedures as a function of instructional level will likely assist in the determination of which generality programming techniques should be used with particular students throughout the course of intervention.

Finally, although the performance feedback interventions resulted in substantial improvements in students' writing fluency over time, a substantial percentage of students

continued to write within the frustrational range at the final intervention session. This suggests that the intervention procedures failed to incite clinically meaningful changes in many students' writing outcomes. One potential reason for this is that some students may have received consistently negative feedback regarding their performance, which could have served as a punisher. In addition, it is unclear to what extent (a) positive feedback serves as a reinforcer and (b) negative feedback serves as a prompt to change behavior. Thus, future research should examine individual-level data to answer these important questions.

Conclusions

Despite the need for empirically-validated interventions to improve elementary-aged students' basic writing skills (National Commission on Writing, 2003), little research in the fields of education and school psychology has systematically examined instructional practices designed to develop students' writing fluency over time. Although performance feedback is one method that has been shown to positively affect students' writing skills, previous research suggests this procedure may result in short-term gains with limited evidence of generalization (Hier & Eckert, 2014). The purpose of this study was to examine the effectiveness of explicitly programming generality into a performance feedback intervention on students' generalized writing outcomes.

Overall, results of this study support previous research (e.g., Eckert et al., 2006) that suggests providing elementary-aged students with weekly performance feedback improves writing fluency to a greater extent than engaging in writing practice alone. The addition of generality programming in the form of multiple exemplar training had no effect on students' generalization of fluency gains across stimuli or responses; however, the generality programming procedures did positively impact students' maintenance of intervention effects. To improve

generalized effects on writing fluency, future research should strive to systematically manipulate

instructional practices as a function of student responding.

Table 1

Student Demographic Information (N = 118)

	Total Sample		Practice		Feedback		Generality			
				Condition						
Characteristics	%	(*n*)	%	(*n*)	%	(*n*)	%	(*n*)	χ^2	*p*
Sex									1.57	.46
Female	42.4	(50)	47.4	(18)	45.9	(17)	34.9	(15)		
Male	57.6	(68)	52.6	(20)	54.1	(20)	65.1	(28)		
Ethnicity									--	--
Hispanic or Latino	0.0	(0)	0.0	(0)	0.0	(0)	0.0	(0)		
Not Hispanic or Latino	100	(118)	100	(38)	100	(37)	100	(43)		
Race									6.18	.40
American Indian or Alaska Native	0.8	(1)	0.0	(0)	0.0	(0)	2.3	(1)		
Asian	0.8	(1)	0.0	(0)	2.7	(1)	0.0	(0)		
Black or African American	0.8	(1)	0.0	(0)	2.7	(1)	0.0	(0)		
White	97.5	(115)	100	(38)	94.6	(35)	97.7	(42)		
Special Education Eligibility									--	--
General Education	100	(118)	100	(38)	100	(37)	100	(43)		
Special Education	0.0	(0)	0.0	(0)	0.0	(0)	0.0	(0)		
	M	(*SD*)	*M*	(*SD*)	*M*	(*SD*)	*M*	(*SD*)	*F*	*p*
Age	8.50	(0.49)	8.51	(0.48)	8.54	(0.48)	8.46	(0.52)	0.27	.76

Table 2

CBM-WE Story Starter Prompts

1. I was on my way home from school and…[a]

2. I found a note under my pillow that said…[b]

3. I was talking to my friends when all of a sudden…[c]

4. I once had a magic pencil and…[c]

5. I was chewing a piece of bubble gum when...[c]

6. One day I went for an airplane ride and...[c]

7. I was playing outside when a spaceship landed and...[c]

8. I opened the front door very carefully and...[c]

9. I was sleeping soundly when…[b]

10. One night I had a strange dream about…[d]

Notes. [a]Used during the baseline session. [b]Used during stimulus generalization sessions. [c]Used during intervention sessions. [d]Used during the maintenance session.

Table 3

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 Sequences

Table 4

Response Generalization and Generalization Training Probe Essay Topics

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

2. Write a composition about how birds are the same as and different from butterflies.[b]

3. Write a composition about how police officers are the same as and different from fire

fighters.[b]

4. Write a composition about how summer is the same as and different from fall.[b]

5. Write a composition about how school is the same as and different from home.[a]

Notes. [a]Used for response generalization measure. [b]Used for generality training probes during intervention sessions.

Table 5

Descriptive Statistics for Procedural Integrity Assessments

Phase/Condition	Sessions Assessed			Percentage of Steps Completed		
	%	(*n*)		*M*	*SD*	Range
Eligibility	85.7	6				
Baseline	100	7		100	0	--
Practice-Only	61.1	11		100	0	--
Performance Feedback	55.6	10		100	0	--
Generality Programming	61.1	11		100	0	--
Stimulus Generalization	71.4	5		100	0	--
Response Generalization	85.7	6		100	0	--
Maintenance	71.4	5		100	0	--
Overall	68.5	61		100	0	--

Notes. Eligibility procedural integrity assessment contained 34 steps. Baseline procedural integrity assessment contained 29 steps. Practice-only procedural integrity assessment contained between 20 and 40 steps; performance feedback procedural integrity assessment contained between 27 and 43 steps; generality programming procedural integrity assessment contained between 28 and 43 steps; stimulus generalization procedural integrity assessment contained 16 steps; response generalization procedural integrity assessment contained 15 steps; maintenance procedural integrity assessment contained 18 steps.

Table 6

Students' Average Scores on Baseline Measures of Writing Performance

Measure	Practice-Only		Feedback		Generality		F	p
	M	(SD)	M	(SD)	M	(SD)		
Correct Writing Sequences[a]	21.29	(9.50)	23.08	(10.38)	24.84	(12.23)	1.08	.34
Total Words Written[a]	24.37	(9.69)	26.73	(10.31)	26.86	(11.03)	0.71	.50
Essay Composition[b]	104.94	(11.18)	104.60	(9.55)	106.44	(11.81)	0.29	.75
Grammar and Mechanics[b]	106.97	(16.36)	102.31	(17.98)	107.67	(16.37)	1.04	.36
Paragraph Copying[c]	21.84	(6.78)	21.78	(5.79)	20.98	(6.91)	0.23	.80

Notes. [a]Metric obtained from the baseline Curriculum-Based Measurement in Written Expression probe. [b]Standard scores on the Essay Composition subtest of the WIAT-III with $M = 100$ and $SD = 15$. [c]Measured by number of correctly copied words on the Paragraph Copying Task.

Table 7

Descriptive Statistics and Correlations between Initial Measures of Writing Performance

Measure	Total Sample		1	2	3	4	5
	M	*(SD)*					
1. Paragraph Copying	21.51	(6.49)	--				
2. Essay Composition[a]	105.34	(10.82)	.42*	--			
3. Grammar and Mechanics[a]	105.66	(16.93)	.51*	.66*	--		
4. CWS[b]	23.14	(10.84)	.55*	.46*	.58*	--	
5. TWW[b]	26.02	(10.36)	.48*	.33*	.32*	.87*	--

Note. TWW = Total Words Written, CWS = Correct Writing Sequences.

[a]Standard scores on the Essay Composition subtest of the WIAT-III with *M* =100 and *SD* = 15.

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

*p< .01.

Table 8

Teachers' Mean Scores on the Writing Orientation Scale

Factor	M	SD
Correct Writing[a]	3.40	1.26
Explicit Instruction[b]	5.00	0.86
Natural Learning[c]	4.05	1.22

Notes. $N = 6$. Answers were based on a Likert-type scale where 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 9

Ratings of Teachers' Instructional Practices

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?	0	0	29%	57%	14%	0	0
2. How often do you re-teach writing skills and strategies?	0	14%	14%	57%	14%	0	0
3. How often do you conference with students about their writing?	0	29%	43%	29%	0	0	0
4. How often do students share their writing with their peers?	0	29%	57%	0	14%	0	0
5. How often do students help each other with their own writing?	0	43%	29%	14%	0	14%	0
6. How often do students select their own writing topics?	0	14%	29%	43%	14%	0	0
7. How often do students use invented spelling in their writing?	0	0	0	29%	14%	29%	29%
8. How often do you specifically teach handwriting skills?	0	0	0	29%	57%	14%	0
9. How often do you specifically teach spelling skills?	0	0	0	50%	33%	17%	0
10. How often do you specifically teach grammar skills?	0	0	29%	14%	57%	0	0
11. How often do you specifically teach planning and revising strategies in writing?	0	29%	14%	29%	14%	14%	0

Notes. $N = 7$. One teacher did not respond to Item 9.

Table 10

Parameter Estimates and Significance Tests from Multilevel Models

Model	Effect	Estimate	*SE*	df	*t*	*p*
Practice vs. Feedback	Intercept of Control	23.20	1.50	72	15.51	< .001
	Group Difference at Baseline	0.62	2.09	350	0.30	.77
	Slope of Control	0.90	0.29	73	3.07	.003
	Group Difference in Slopes	0.99	0.42	350	2.36	.02
Practice vs. Generality	Intercept of Control	22.68	1.50	78	15.14	< .001
	Group Difference at Baseline	-1.64	2.03	381	-0.81	.42
	Slope of Control	0.90	0.28	79	3.20	.002
	Group Difference in Slopes	0.74	0.39	381	1.92	.056
Feedback vs. Generality	Intercept of Feedback	22.93	1.46	77	15.69	< .001
	Group Difference at Baseline	-2.23	1.94	373	-1.15	.25
	Slope of Feedback	1.88	0.29	78	6.56	< .001
	Group Difference in Slopes	-0.23	0.39	373	-0.60	.55

Table 11

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

| | Practice-Only | | | | Performance Feedback | | | | Generality Programming | | | | ANCOVA | |
| | Baseline | | Post-Intervention | | Baseline | | Post-Intervention | | Baseline | | Post-Intervention | | | |
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	F (1, 109)	Partial η^2
CWS	29.14	12.85	27.53	13.77	30.97	13.79	31.34	13.03	30.38	14.65	32.94	14.43	2.96*	.05

Note. CWS = Correct Writing Sequences.

*p = .03 (based on one-tailed test).

Table 12

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

| | Practice-Only | | | | Performance Feedback | | | | Generality Programming | | | | ANCOVA | |
| | Baseline | | Post-Intervention | | Baseline | | Post-Intervention | | Baseline | | Post-Intervention | | | |
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	F (1, 111)	Partial η^2
CWS	55.14	35.88	69.74	39.96	50.94	30.34	84.58	44.55	47.98	25.08	84.20	38.91	2.44*	.04

Note. CWS = Correct Writing Sequences.

**p = .95 (based on one-tailed test).

Table 13

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

	Practice-Only				Performance Feedback				Generality Programming			
	M	*SD*	Mean Percent Change	Percent Maintained	*M*	*SD*	Mean Percent Change	Percent Maintained	*M*	*SD*	Mean Percent Change	Percent Maintained
Final Intervention Session	30.91	15.58	--	--	36.75	14.30	--	--	34.38	13.84	--	--
Maintenance Session	32.91	13.63	23.85	54.5	33.22	14.69	-7.38	35.3	38.08	15.51	18.47	63.4

Notes. Mean percent change was calculated by averaging the percent change score for each individual. Percent maintained represents

the percentage of students in each condition who obtained a percent change score of 0 or greater.

Table 14

Shifts in Instructional Level from Baseline to the Final Intervention Session

Practice-Only Condition	Baseline		Final Intervention		Maintenance	
	%	(*n*)	%	(*n*)	%	(*n*)
Frustrational	92.1	(35)	57.9	(22)	45.5	(15)
Instructional	0.0	(0)	7.9	(3)	18.2	(6)
Mastery	7.9	(3)	34.2	(13)	36.4	(12)

Performance Feedback Condition	Baseline		Final Intervention		Maintenance	
	%	(*n*)	%	(*n*)	%	(*n*)
Frustrational	86.5	(32)	35.1	(13)	47.1	(16)
Instructional	8.1	(3)	2.7	(1)	8.8	(3)
Mastery	5.4	(2)	62.2	(23)	44.1	(15)

Generality Programming Condition	Baseline		Final Intervention		Maintenance	
	%	(*n*)	%	(*n*)	%	(*n*)
Frustrational	83.7	(36)	46.5	(20)	39.0	(16)
Instructional	9.3	(4)	14.0	(6)	9.8	(4)
Mastery	7.0	(3)	39.5	(17)	51.2	(21)

Total Sample	Baseline		Final Intervention		Maintenance	
	%	(*n*)	%	(*n*)	%	(*n*)
Frustrational	87.3	(103)	46.6	(55)	43.5	(47)
Instructional	5.9	(7)	8.5	(10)	12.0	(13)
Mastery	6.8	(8)	44.9	(53)	44.4	(48)

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 15

Students' Intervention Acceptability Ratings

	Total Sample[b]		Practice-Only[c]		Feedback[d]		Generality Programming[e]	
Procedures associated with CBM-WE	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*
How much do you like writing stories with us each week?	3.91	(1.35)	3.84	(1.46)	3.82	(1.40)	4.05	(1.20)
How much do you like being told what to write about?	3.25	(1.63)	3.00	(1.63)	3.15	(1.56)	3.56	(1.66)
Were there times when you didn't want to write stories with us?[a]	3.79	(1.45)	3.32	(1.58)	3.94	(1.50)	4.10	(1.18)
Were there any times when you wished you could work more on writing stories with us?	3.65	(1.57)	3.30	(1.60)	3.62	(1.60)	4.00	(1.48)
Procedures associated with performance feedback					*M*	*(SD)*	*M*	*(SD)*
How much do you like being told how many words you wrote?					4.39	(1.30)	4.76	(0.70)
How much do you think it helps you when you were told how many words you wrote?					4.24	(1.35)	4.41	(1.02)
Procedures associated with practice	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*
Do you think your writing has improved?	3.98	(1.27)	3.81	(1.41)	4.12	(1.27)	4.02	(1.13)
Do you think your writing has gotten worse? [a]	4.56	(0.98)	4.62	(0.89)	4.65	(0.92)	4.44	(1.10)
Overall acceptability			3.65	(0.97)	4.00	(0.89)	4.17	(0.86)

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]$n = 112$. [c]$n = 37$. [d]$n = 34$. [e]$n = 41$.

Table 16

Teachers' Intervention Acceptability Ratings

Item	M	(SD)
This would be an acceptable intervention for students' writing difficulties.	5.67	(0.52)
Most teachers would find this intervention appropriate for writing difficulties in addition to the one described.	5.20	(0.45)
This intervention should prove effective in changing students' writing difficulties.	5.17	(0.41)
I would suggest the use of this intervention to other teachers.	5.50	(0.55)
The students' writing difficulties are severe enough to warrant the use of this intervention.	4.83	(0.41)
Most teachers would find this intervention suitable for the writing difficulties described.	5.17	(0.41)
I would be willing to use this intervention in my classroom.	5.67	(0.52)
This intervention would not result in negative side effects for the students.	5.67	(0.52)
This intervention would be appropriate for a variety of students.	5.50	(0.55)
This intervention is consistent with those I have used in school.	4.33	(1.63)
The intervention is a fair way to handle the students' writing difficulties.	5.17	(0.41)
This intervention is reasonable for the writing difficulties described.	5.00	(0.00)
I like the procedures used in this intervention.	5.67	(0.52)
This intervention is a good way to handle the students' writing difficulties.	5.33	(0.52)
Overall, this intervention would be beneficial for the students.	5.67	(0.52)
Overall acceptability	5.31	(0.38)

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

Planning		→	Translating	→	Reviewing	

Idea generation	Organizing	Goal-setting		Evaluation	Revision
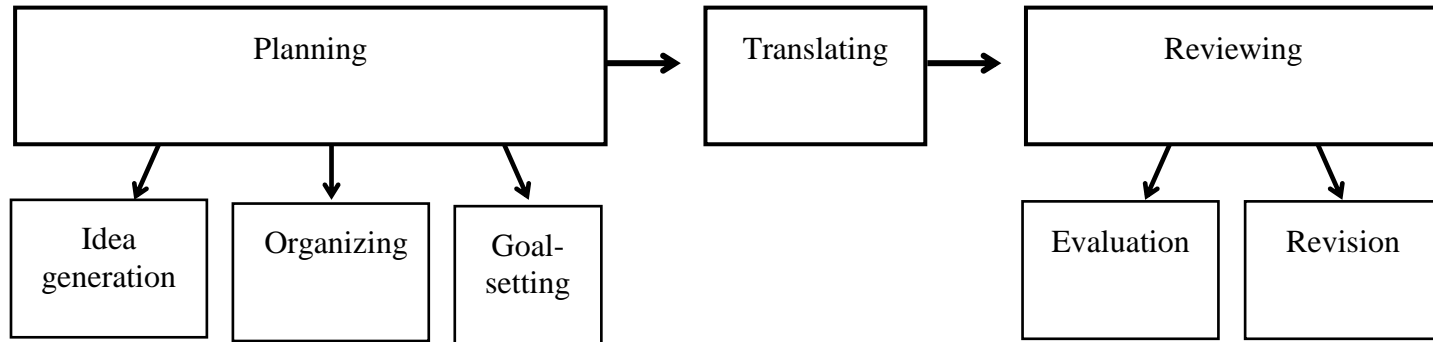
Figure 2. Berninger and colleagues' (1992) Component Processes of Writing

```
┌─────────────────────────────────────┐
│              Translating             │
└─────────────────────────────────────┘
        │                   │
        ▼                   ▼
┌──────────────────┐ ┌──────────────────┐
│ Text generation  │ │  Transcription   │
└──────────────────┘ └──────────────────┘
                              │
                              ▼
                    ┌──────────────────┐
                    │   Handwriting,   │
                    │     Spelling,    │
                    │     Fluency      │
                    └──────────────────┘
```
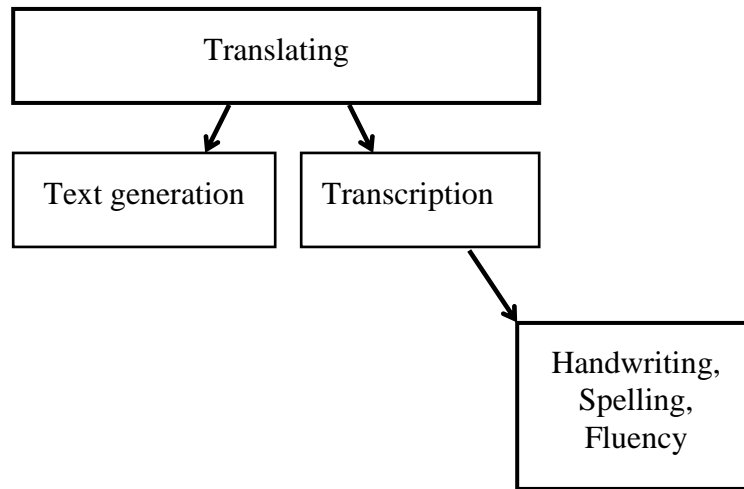
Figure 3. Participant flow chart following Consolidated Standards of Reporting Trials guidelines
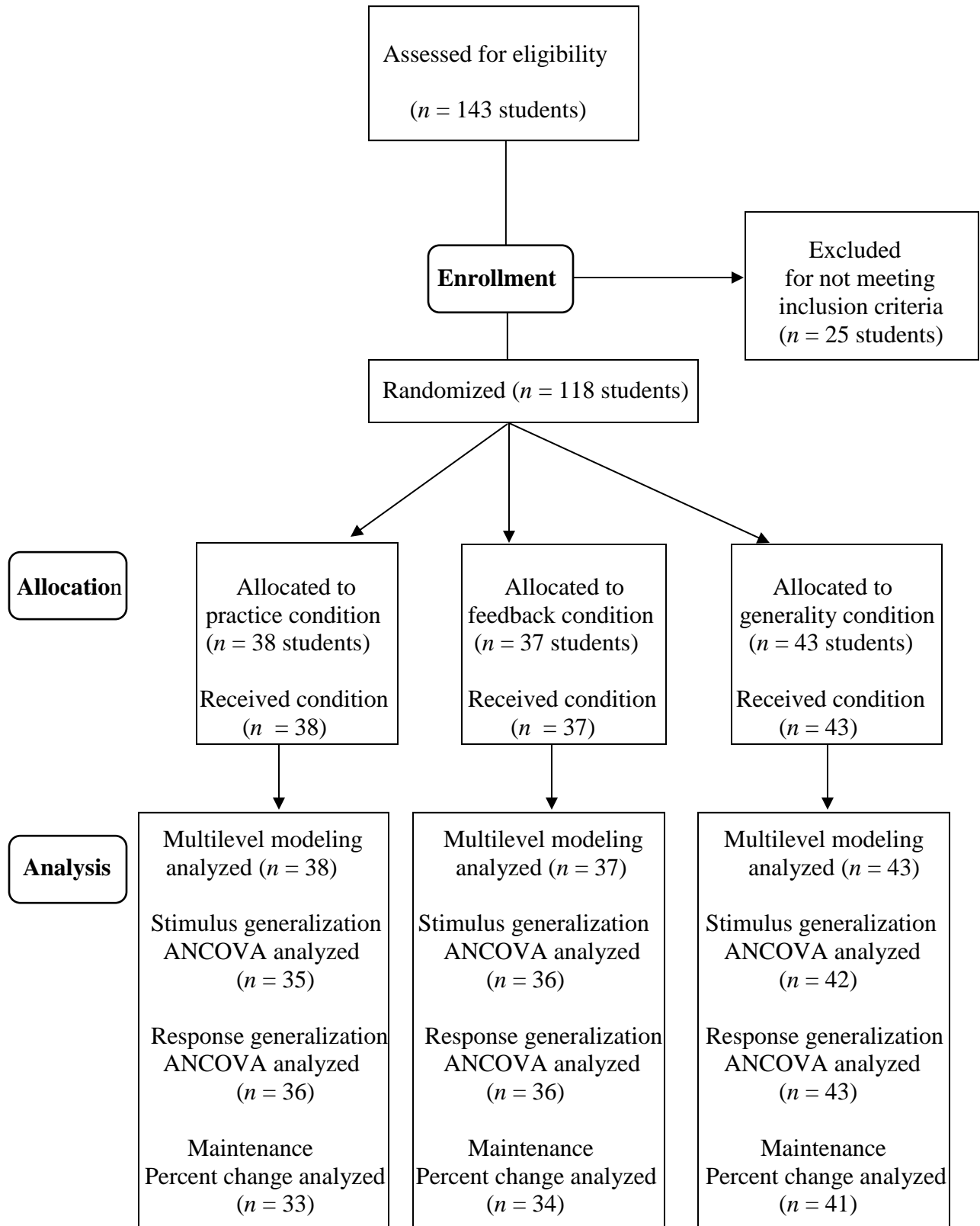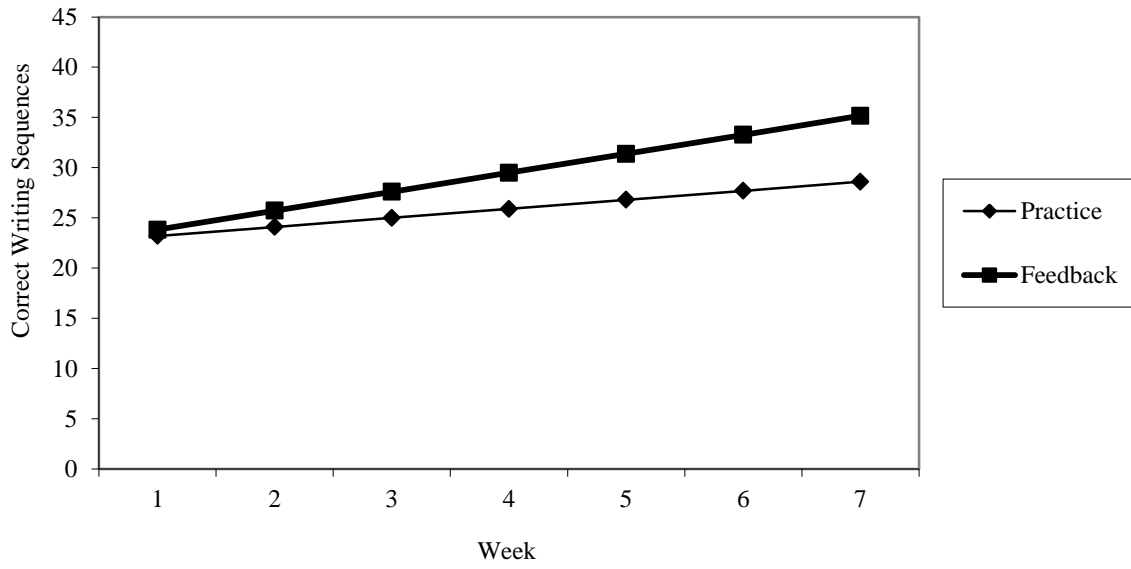
Figure 4. Growth trajectories by condition (i.e., practice-only and performance feedback),

reflecting students' average gains of correct writing sequences.
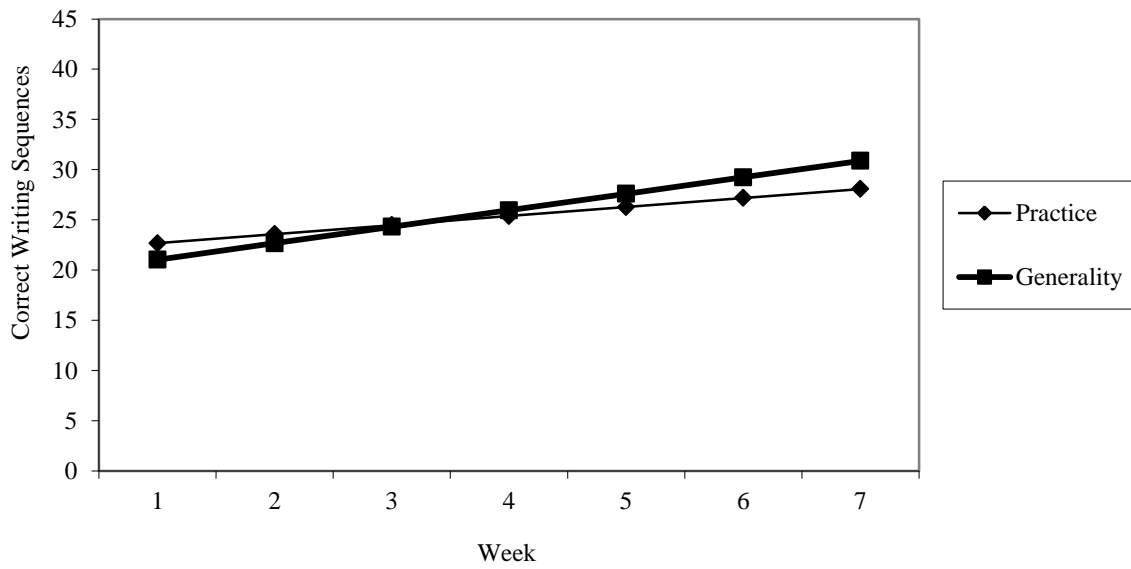
Figure 5. Growth trajectories by condition (i.e., practice-only and generality programming), reflecting students' average gains of correct writing sequences.
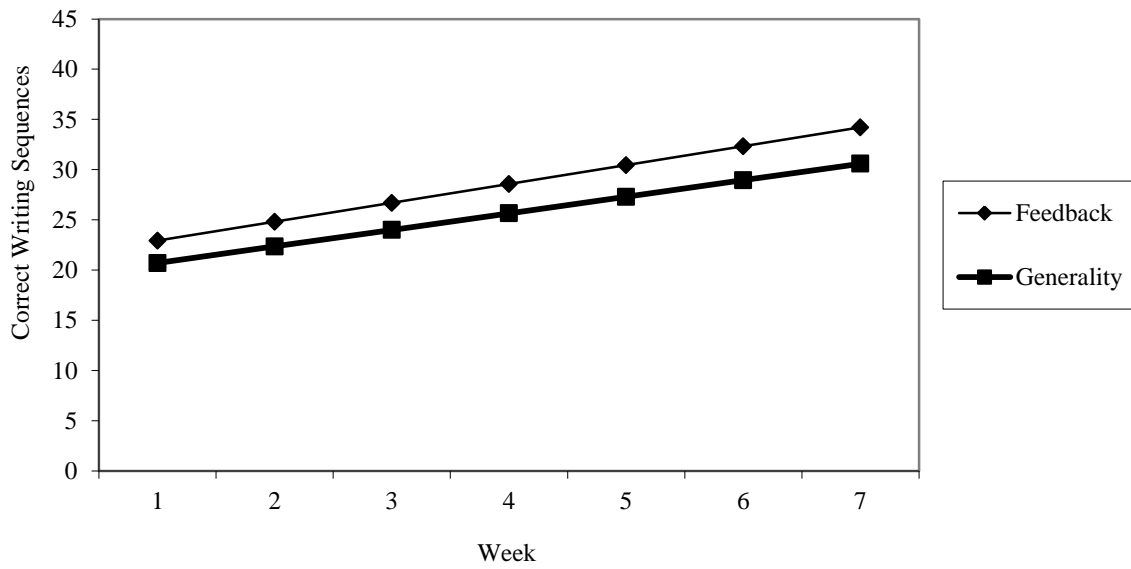
Figure 6. Growth trajectories by condition (i.e., performance feedback and generality

programming), reflecting students' average gains of correct writing sequences.

List of Appendices

Appendix A: Parental Informational Letter

Appendix B: Student Assent

Appendix C: Writing packet: Page 1, Identification Information

Appendix D: Writing packet: Page 2, Stop Sign

Appendix E: Writing packet: Story Starter Page with Stop Sign

Appendix F: Writing packet: Story Starter with Writing Lines

Appendix G: Response Generalization Task

Appendix H: Procedural Script for Response Generalization Task

Appendix I: Paragraph Copying Task

Appendix J: Handwriting Proficiency Screening Measure

Appendix K: Student Intervention Acceptability Packet

Appendix L: Intervention Rating Profile-15

Appendix M: Teacher Questionnaire

Appendix N: Schematic of Intervention Procedures

Appendix O: Procedural Script for Individualized Performance Feedback Conditions

Appendix P: Feedback Page for Performance Feedback Conditions

Appendix Q: Procedural Script for Practice-Only Condition

Appendix R: CBM-WE Scoring Manual

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!

Student Assent

Important Question

I would like to work with you each week for the next few 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*

Name:_____

Writing Packet: Page 1, Identification Information

Syracuse University

2013-2014 Writing Project

_____ School

3rd grade

Name: _____

Classroom: _____

Appendix D

Writing Packet: Page 2, Stop Sign
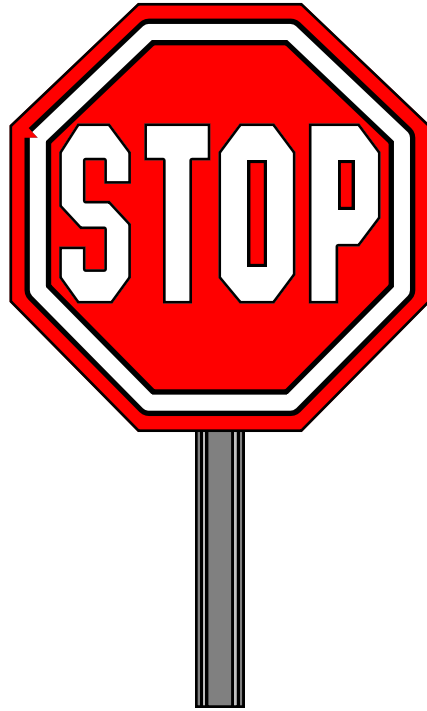
Writing Packet: Story Starter Page with Stop Sign

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

Writing Packet: Story Starter with Writing Lines

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

_____

_____

_____

_____

_____

_____

_____

_____

Keep going ➡
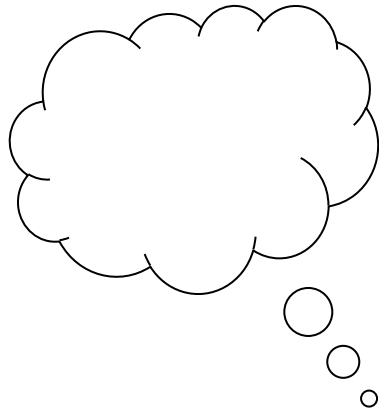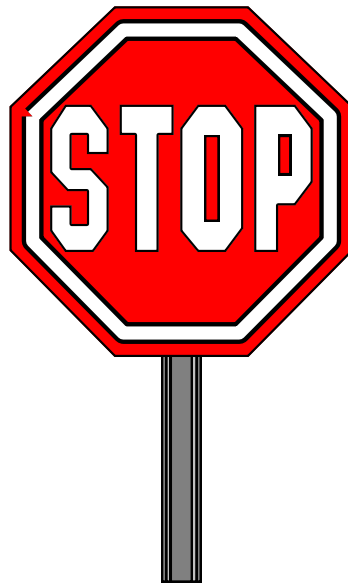
Appendix G

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 the 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:			
Name of secondary research assistant:			
Classroom:			
Date: 6/14/13			
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.	Gain the attention of the class by announcing (in a voice LOUD ENOUGH for all students to hear) that you are ready to begin. Once all students are paying attention, state to the students: "*Hi everyone. If you have not already done so, please clean off the top of your desk, except for a pencil. Two of your classmates will be passing out your packets. Please leave the packets closed until you are told to open them.*"	
2.	State to the students: "*When you get a packet, please do not turn to any of the pages in the packet. Keep the front page with the pencil facing you.*" [The research assistant should point to the pencil on the coversheet of the packet.]	
3.	The research assistant should make sure each student has a pencil and provide pencils to those students without.	
4.	Research assistant should give the packets to two reliable students in the classroom to hand out to their peers.	
5.	After all of the packets have been distributed, state to the students: "*Now, everyone please turn to the first page in your packet that looks like this.*" [Hold up the first page of the Response Generalization measure for students to see].	

6.	The research assistant should make sure all the students are on the correct page.	
7.	State to the students: "*Look at the writing prompt on this page.*" [<u>Hold up the packet and point to the writing prompt</u>.] "*It says, 'Write a composition about how school is the same as and different from home.' This page 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.*"	
8.	Start the stopwatch and time the students for 3 minutes. During this time, monitor students to ensure they are following directions from the previous step, 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. If any student asks what they are supposed to be writing about, point to (but do <u>not</u> orally state) the writing prompt.	
9.	At the end of 3 minutes, state to the students: "*Please raise your pencil high in the air and turn to the next page with lines on it.*"	
10.	Monitor to ensure all students have turned to the correct page and have not started writing yet.	
11.	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.	

	If any student asks what they are supposed to be writing about, point to (but do <u>not</u> orally state) the writing prompt.	
12.	State to the students: *"Okay, you may begin writing now."*	
13.	Discretely start the stopwatch and time the students for 10 minutes. During this time, monitor to ensure students are following directions. Do <u>not</u> prompt students who stop writing early.	
14.	At the end of 10 minutes, state to the students: *"Please stop writing and close your packet. That's all the writing we are going to do today. Thank you for your hard work!"*	
15.	Monitor to ensure all students have stopped writing and collect the packets.	
Total number of steps completed:		

Appendix I

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.

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. _____
- - - - - - - -
_____

Student Intervention Acceptability Packet

Question #1

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

| Not at all | A little bit | Some | A lot | Very, very much |
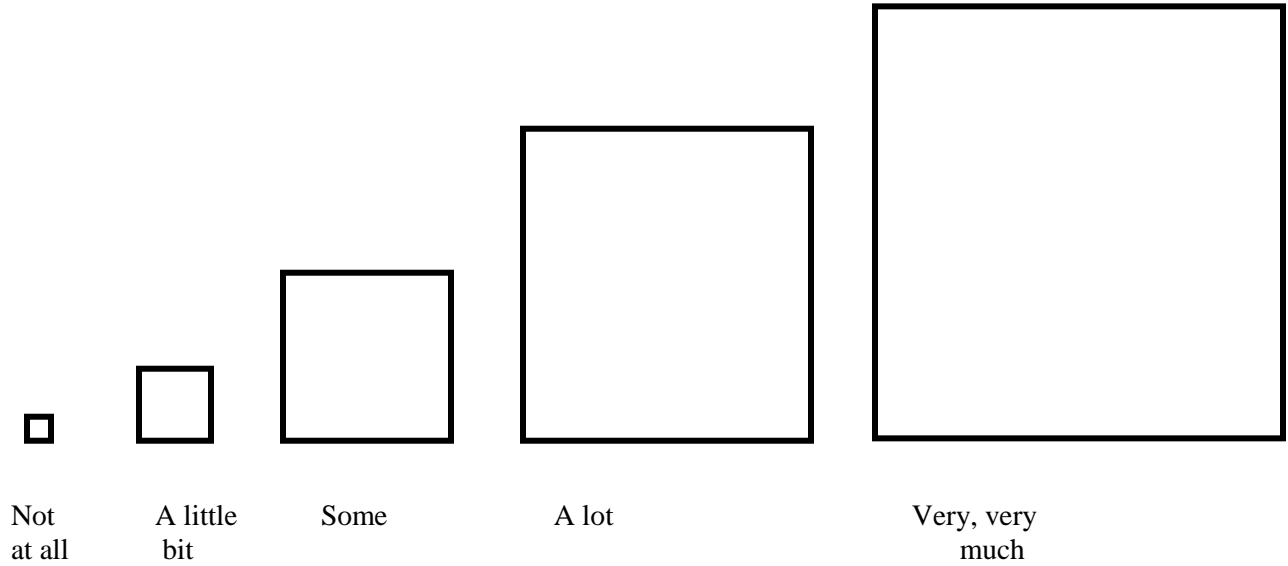
Question #2

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

| Not at all | A little bit | Some | A lot | Very, very much |
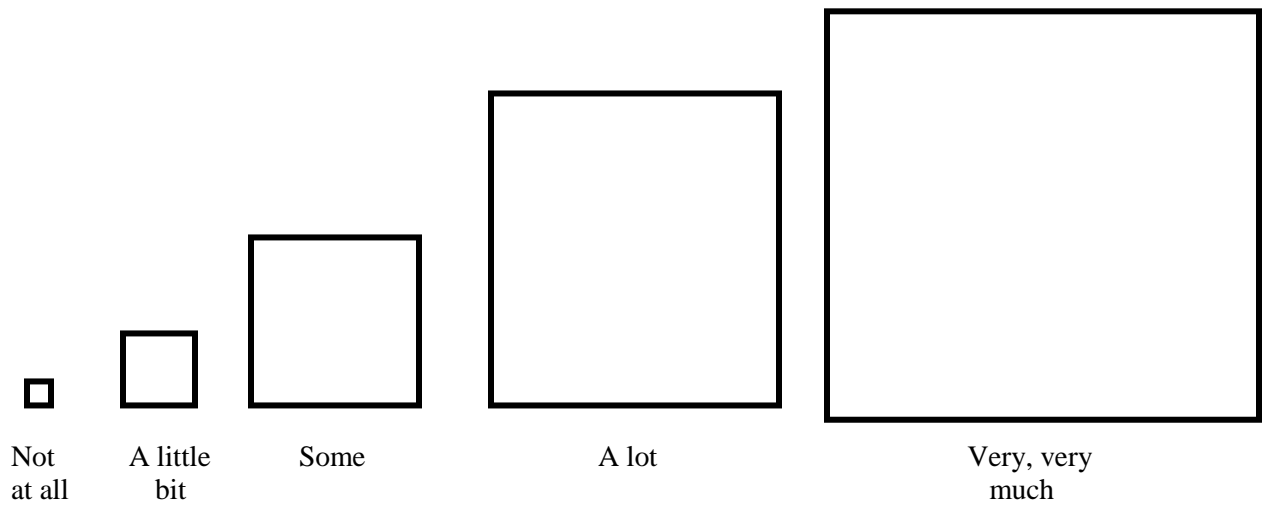
Question #3
Were there times when you didn't want to write stories with us?

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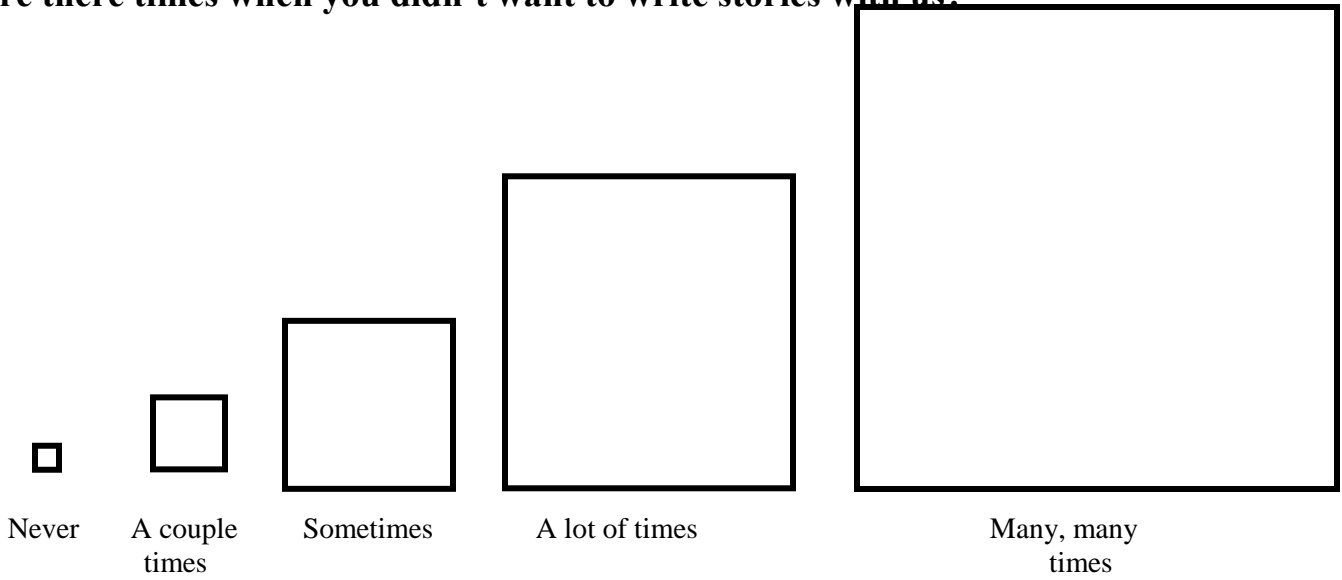
times   times

Question #4
Were there any times when you wished you could work more on writing stories with us?

Never   A couple   Sometimes   A lot of times   Many, many

times   times
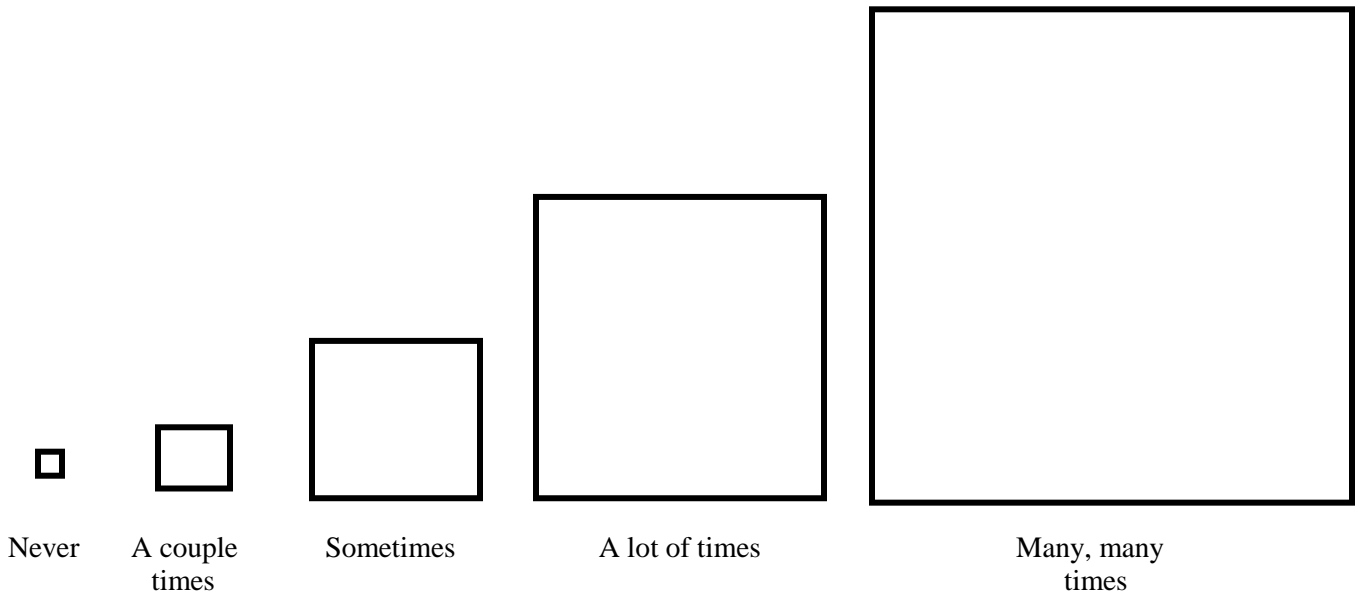
Question #5
How much do you like being told how many words you wrote?

Not at all	A little bit	Some	A lot	Very, very much
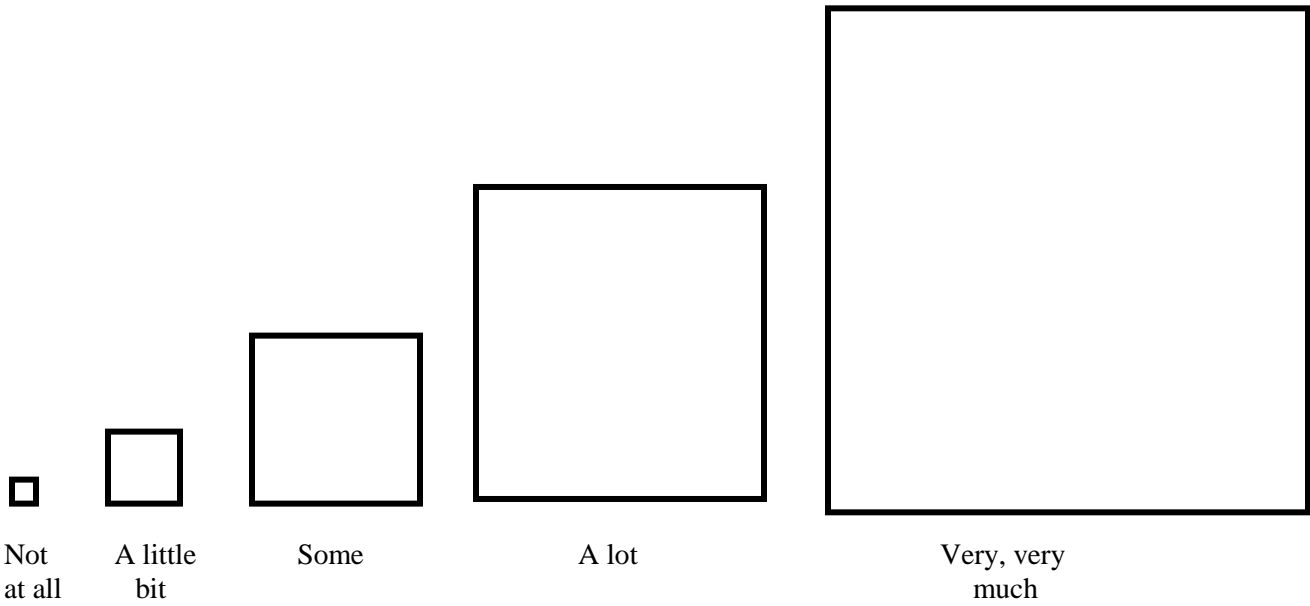
Question #6

How much do you think it helped you when you were told how many words you wrote?

Not at all	A little bit	Some	A lot	Very, very much
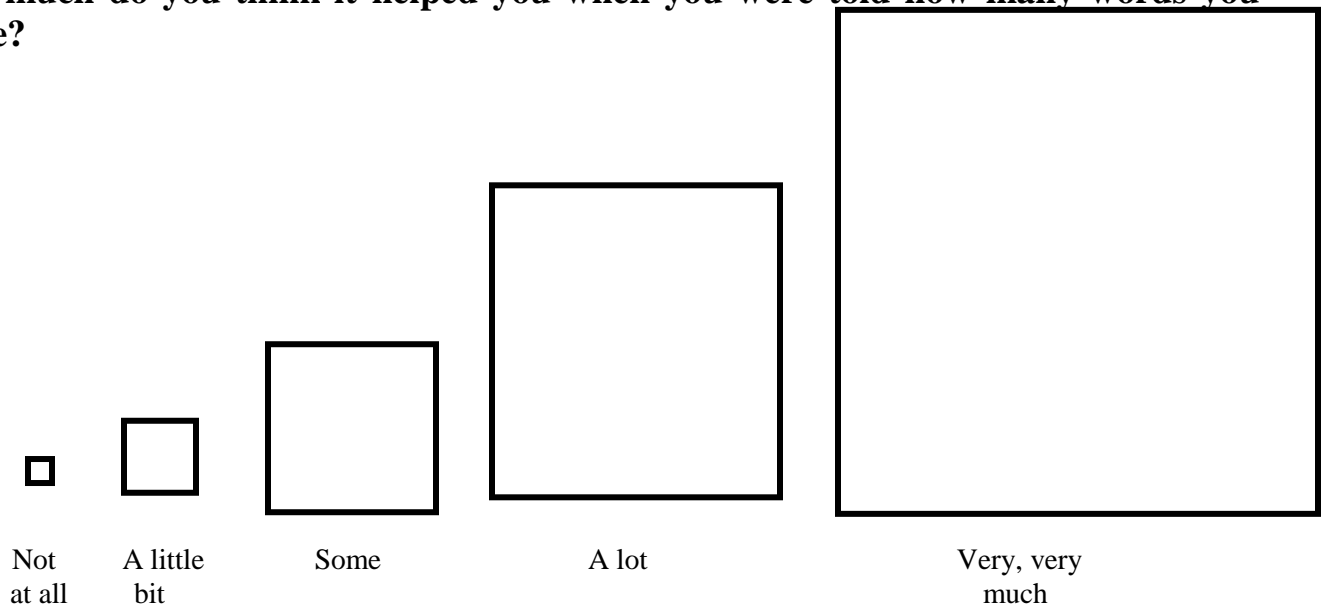
Question #7
Do you think your writing has improved?

Not at all	A little bit	Some	A lot	Very, very much
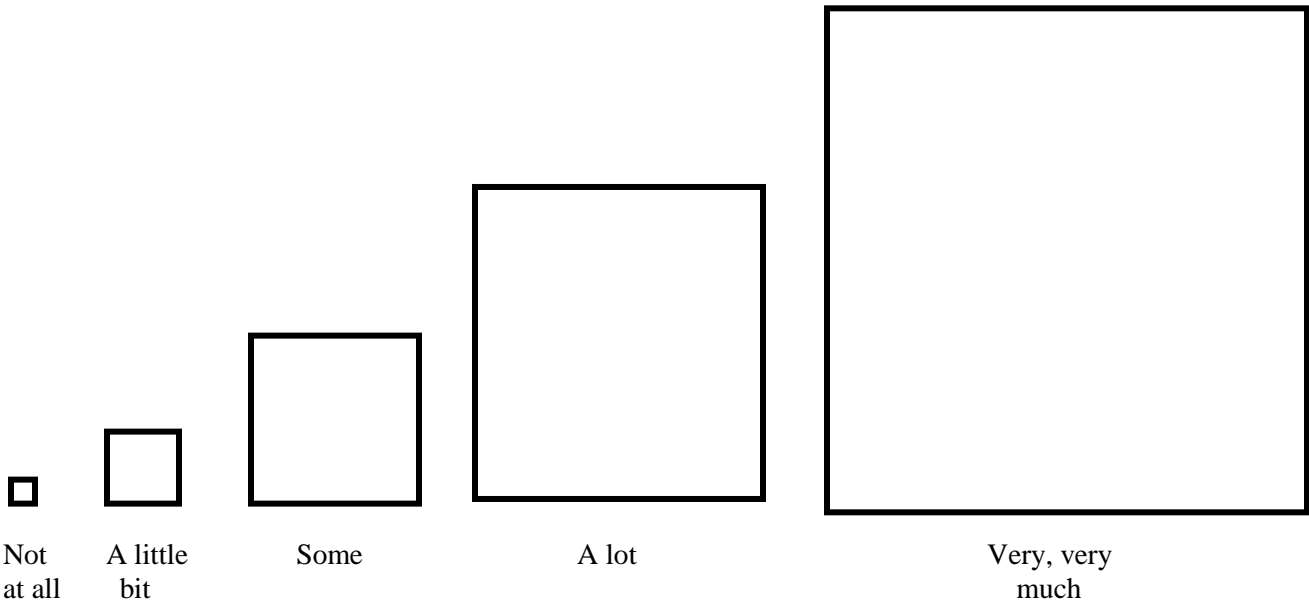
Question #8
Do you think your writing has gotten worse?

Not at all	A little bit	Some	A lot	Very, very much
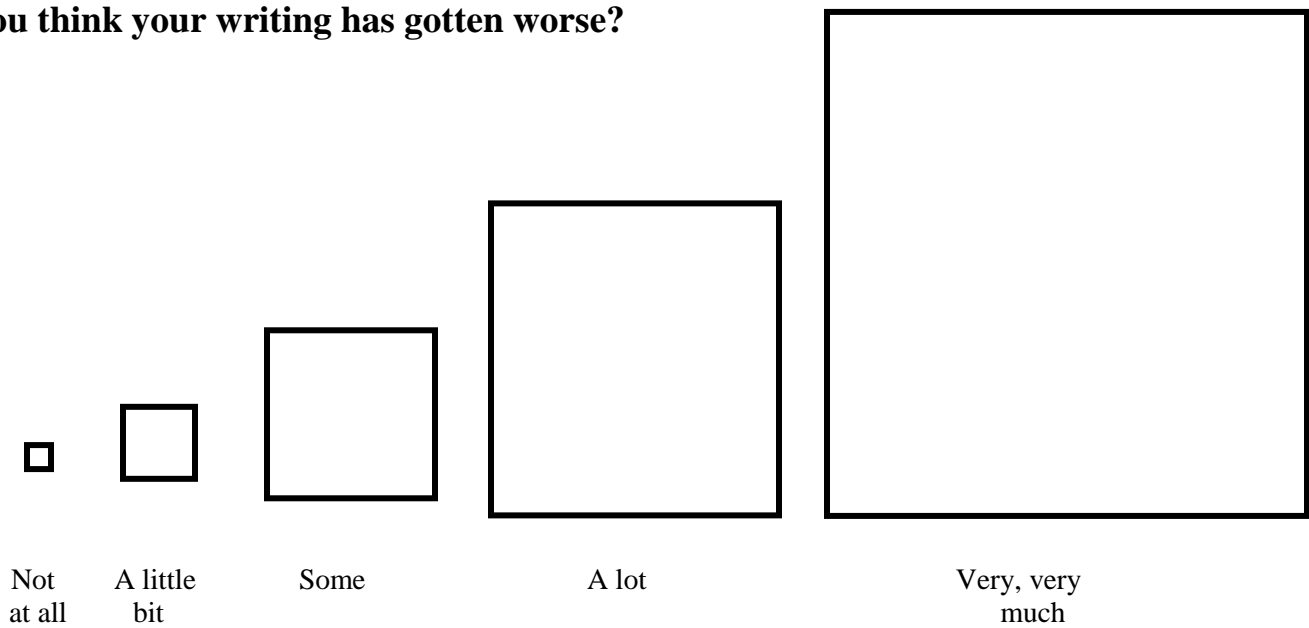
Intervention Rating Profile – 15 (IRP-15) – Teacher Version

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

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 to 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 spelling.	1	2	3	4	5	6
9	Formal instruction in writing is necessary to insure adequate development of all 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 o 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 N

Schematic of Intervention Procedures

Random assignment of students to one of three conditions:		
Practice-Only	**Performance Feedback**	**Generality Programming**
--	Provided with individualized performance feedback	Provided with individualized performance feedback
Administered CBM-WE story-starter	Administered CBM-WE story-starter	Administered CBM-WE story-starter or abbreviated response generalization probe (alternates weekly)
Given 1 minute to plan the composition	Given 1 minute to plan the composition	Given 1 minute to plan the composition
Given 3 minutes to write the composition	Given 3 minutes to write the composition	Given 3 minutes to write the composition

Appendix O

Procedural Script for Individualized Performance Feedback Conditions

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

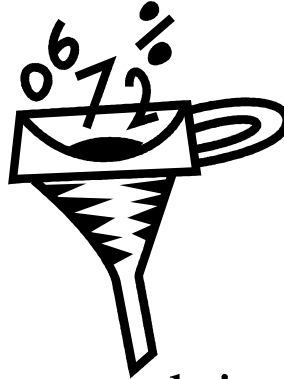| I.	Identifying Information			
Name of primary research assistant:				
Name of secondary research assistant:		or N/A		
School/Classroom:				
Date:				
Notes:				
II.	Data Collection – Material Preparation	Circle		
a.	Five (5) sharpened pencils		Yes	No
b.	Assessment packets		Yes	No
c.	Experimenter's copy of packet		Yes	No
d.	Two (2) stopwatches		Yes	No
Notes:				
III.	Data Collection Procedures			
[Please check [✔] each box as you complete each step]				✔
1.	Gain the attention of the class by announcing (in a voice LOUD ENOUGH for all students to hear) that you are ready to begin. Once all students are paying attention, state to the students: "*Hi everyone. If you have not already done so, please clean off the top of your desk, except for a pencil. Two of your classmates will be passing out your packets. Please leave the packets closed until you are told to open them.* "			
2	Research assistant should give the packets to reliable students in the classroom to hand out to their peers.			
3.	After all of the packets have been distributed, state to the students: "*Everyone has a symbol at the top of the front page of your packet.*" [Point to the symbol on the examiner copy so the students can see where to look] "*When I call your group symbol, please line up at the door with your packet and your pencil.*" "*The Frog Group will be staying in this classroom to work with us.* "*The Horse Group will be will be going to Mrs. _____'s classroom. Please line up now and show me how you walk quietly through the halls.*" "*The Lion Group will be will be going to Mrs. _____'s classroom. Please line up now and show me how you walk quietly through the halls.*"			

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. If 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 sheet of paper that says the correct 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 Frog Group. Please turn to the 2nd page of your packet that has three faces at the top of the page. Please watch me while I explain the directions. Each face is for a person who is excited to write for this activity* [point to the excited face]*, doesn't mind writing for this activity* [point to the straight face]*, or is not excited to write for this activity* [point to the sad face]. Circle the face that best describes how you feel about writing with us today.*"	
6.	Once the students have finished, state: "*Now turn to the next page of your packet that has a 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.*"	
7.	The research assistant should scan the room to make sure all of the students are on the correct page.	
8.	State to the students: "*The box in the middle of the page* [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, that means 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.* *If you have an equal sign instead of an arrow, that means you wrote the same number of words as you did the last time I worked with you.*"	

|		
	Every week when we work with you, we are going to tell you how you are doing with your writing."	
9.	The research assistant should monitor the students for questions.	
10.	State to the students: *"Everyone should still be on the page that has your number on it. 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."*	
11.	State to the students: *"Please turn to the next page of your packet. This page has a thought bubble at the top of the page.*"	
12.	State to the students: *"For the next minute think about writing a story that begins with this sentence –* **I was playing outside when a spaceship landed and. . .** *Remember, take time to plan your story. A well-written story usually has a beginning, a middle, and an 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 yet. Just think of a story that begins with this sentence:* **I was playing outside when a spaceship landed and. . .**"	
13.	The research assistant should begin the stopwatch and time the students for 1 minute.	
14.	The research assistants should monitor the students to ensure they are not writing the story yet.	
15.	After 30 seconds, state to the students: *"You should be thinking about:* **I was playing outside when a spaceship landed and. . .**"	
16.	At the end of 1 minute, state to the students:	

	"Okay, stop thinking, turn to the next page of your packet, and raise your pencil high in the air."	
17.	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."*	
18.	State to the students: *"Okay, you can start writing."* The research assistant should begin the stop watch and time the students for 3 minutes.	
19.	The research assistant should monitor the students and make sure students are following the directions. 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 don't need to copy the words that have been provided."*	
20.	After 1 minute, 30 seconds has elapsed, state to the students: *"You should be writing about* – **I was playing outside when a spaceship landed and. . .***"*	
21.	After 3 minutes has elapsed, state to the students: *"Okay, please close your packets. That is all of the writing that we are going to do today. All of you did a very nice job following my directions. We will come around to collect your packets. Thank you for working with us today."*	
22.	The research assistant should collect all of the packets.	
23.	State to the students: *"All of the students in Mrs. _____ 's (and _____ 's) classroom, please pick up your pencil and line up at the door."*	
24.	The research assistants should then assist the students in getting back to their classrooms quickly and quietly.	
Total number of steps completed:		

Feedback Page for Performance Feedback Conditions

Here is how you are doing in writing:
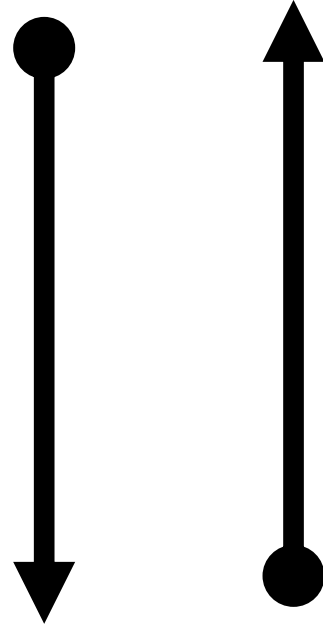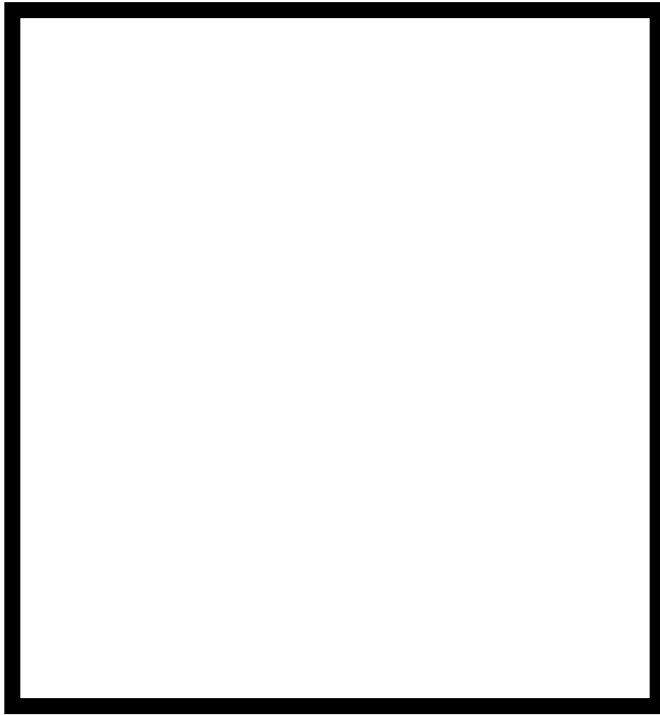
Appendix Q

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
Notes:			
III.	Data Collection Procedures		✔
[Please check [✔] each box as you complete each step]			
1.	Gain the attention of the class by announcing (in a voice LOUD ENOUGH for all students to hear) that you are ready to begin. Once all students are paying attention, state to the students: "*Hi everyone. If you have not already done so, please clean off the top of your desk, except for a pencil. Two of your classmates will be passing out your packets. Please leave the packets closed until you are told to open them.* "		
2	Research assistant should give the packets to reliable students in the classroom to hand out to their peers.		
3.	After all of the packets have been distributed, state to the students: "*Everyone has a symbol at the top of the front page of your packet.*" [Point to the symbol on the examiner copy so the students can see where to look] "*When I call your group symbol, please line up at the door with your packet and your pencil.*" "*The Lion Group will be staying in this classroom to work with us.* "*The Horse Group will be will be going to Mrs. _____ 's classroom. Please line up now and show me how you walk quietly through the halls.*"		

	"The Frog Group will be will be going to Mrs. _____'s classroom. Please line up now and show me how you walk quietly through the halls."	
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. If 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 sheet of paper that says the correct 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 Lion Group. Please turn to the 2nd page of your packet that has three faces at the top of the page. Please watch me while I explain the directions. Each face is for a person who is excited to write for this activity* [<u>point to the excited face</u>]*, doesn't mind writing for this activity* [<u>point to the straight face</u>]*, or is not excited to write for this activity* [<u>point to the sad face</u>]. *Circle the face that best describes <u>how you feel</u> about writing with us today."*	
6.	Once the students have finished, state: *"Now turn to the next page of your packet that has a stop sign in the middle of the page. Today 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."*	
7.	State to the students: *"Please turn to the next page of your packet. This page has a thought bubble at the top of the page."* Research assistants should monitor the students to ensure they are on the correct page.	
8.	State to the students: *"For the next minute think about writing a story that begins with this sentence –* **I was playing outside when a spaceship landed and. . .**	

	Remember, take time to plan your story. A well-written story usually has a beginning, a middle, and an 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 yet. Just think of a story that begins with this sentence:* **I was playing outside when a spaceship landed and. . .**"	
9.	The research assistant should begin the stopwatch and time the students for 1 minute.	
10.	The research assistants should monitor the students to ensure they are not writing the story yet.	
11.	After 30 seconds, state to the students: *"You should be thinking about:* **I was playing outside when a spaceship landed and. . .**"	
12.	At the end of 1 minute, state to the students: *"Okay, stop thinking, turn to the next page of your packet, and raise your pencil high in the air."*	
13.	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."*	
14.	State to the students: *"Okay, you can start writing."* The research assistant should begin the stop watch and time the students for 3 minutes.	
15.	The research assistant should monitor the students and make sure students are following the directions. 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 don't need to copy the words that have been provided."*	

16.	After 1 minute, 30 seconds has elapsed, state to the students: "*You should be writing about* – **I was playing outside when a spaceship landed and. . .**"	
17.	After 3 minutes have elapsed, state to the students: "*Okay, please close your packets. That is all of the writing that we are going to do today. All of you did a very nice job following my directions. We will come around to collect your packets. Thank you for working with us today.*"	
18.	The research assistant should collect all of the packets.	
19.	State to the students: "*All of the students in Mrs. _____'s (and _____'s) classroom, please pick up your pencil and line up at the door.*"	
20.	The research assistants should then assist the students in getting back to their classrooms quickly and quietly.	
Total number of steps completed:		

2012-2013 TRAC RESEARCH PROJECT

RA Training Manual: Administration and Scoring of Curriculum-Based Measurement in Written Expression Probes revised 10/4/2012

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:

> One day I was out sailing. A storm carried me far out to
> sea and wrecked my boat on a desert island. . . .
>
> _____
>
> _____
>
> _____
>
> _____
>
> _____
>
> _____

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 writing sequences (CWS)

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:

I woud drink water from the ocean 07
and I woud eat the fruit off of   08
the trees. Then I woud bilit a   07
house out of trees, and I woud   07
gather firewood to stay warm. I   06
woud try and fix my boat in my   08
spare time.   02

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

2 – Words Spelled Correctly (WSC):

The examiner counts up and records only those words in the writing sample that are spelled correctly. <u>Words are considered separately, not within the context of a sentence</u>. 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 <u>correctly spelled words</u> and <u>total words written</u> is provided below:

```
I woud drink water from the ocean . . . . . 06
and I woud eat the fruit off of . . . . . .     07
the trees.  Then I woud bilit a . . . . . .     05
house out of trees, and I woud . . . . . .      06
gather firewood to stay warm.  I . . . . .      06
woud try and fix my boat in my . . . . . .      07
spare time.  . . . . . . . . . . . . . . .      02
```

This sample is found to contain **39 correctly spelled words** and **45 total words written**.

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. Each sequence should be examined in isolation and credit should be given when the sequence is correct (e.g., "seen the") or marked incorrect when the sequence is not correct (e.g., "could seen"). 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:

Because the period is considered essential punctuation, it is joined with the words before and after it to make 2 correct writing sequences.

Since the first word is correct it is marked as a correct writing sequence.

^It^was^dark^.^Nobody^

could^{X}seen^the^trees^of

^the^{X}forrest^X.

Grammatical or syntactical errors are not counted.

Misspelled words are not counted.
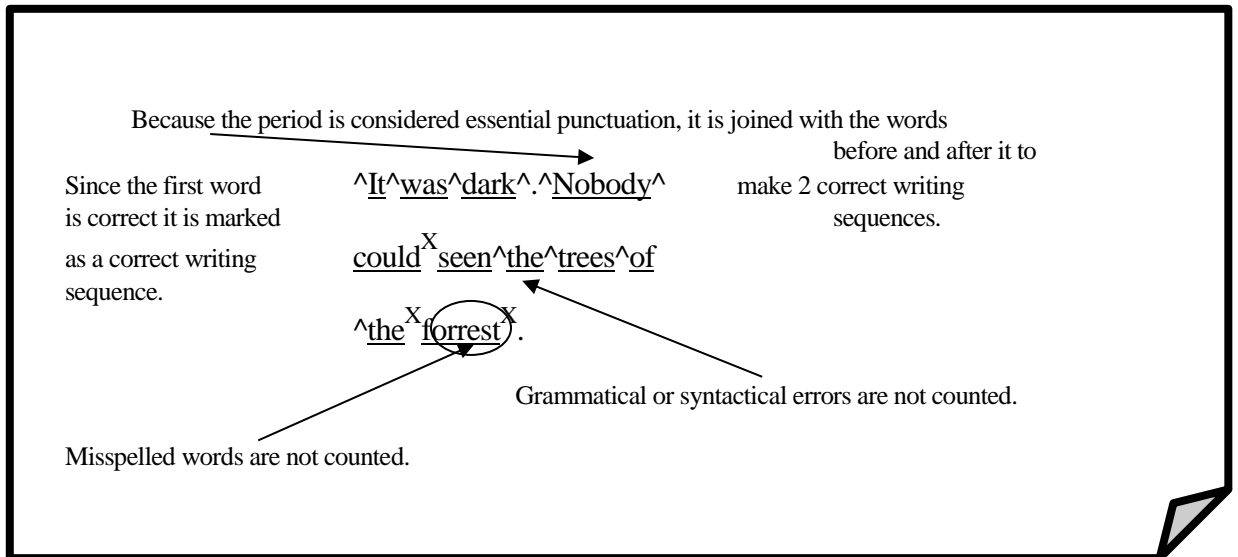
3 – Correct Writing Sequences (CWS):

Rules:

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

Example: ^I^like^the^reb^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^? [note: 'car read' is not syntactically correct]

☑ Semantically correct words make up a correct writing sequence:

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

^Is^that^a^{X}read^{X}car^? [note: 'read car' is not semantically correct]

☑ 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":

  <u>Example:</u> ^The^Terrible^Rotten^Day

☑ For this measure, numerals **are counted**.

  <u>Example:</u> ^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 <u>correct writing sequences</u> is provided below:

<u>^I^{X}woud^{X}drink^water^from^the^ocean</u>	05
<u>^and^I^{X}woud^{X}eat^the^fruit^off^of</u>	06
<u>^the^trees^. ^Then^I^{X}woud^{X}bilit^{X}a</u>	05
<u>^house^out^of^trees,^and^I^{X}woud</u>	06
<u>^{X}gather^firewood^to^stay^warm^.^I</u>	06
<u>^{X}woud^{X}try^and^fix^my^boat^in^my</u>	06
<u>^spare^time^.</u>	03

This sample is found to contain **37 correct writing sequences**, **39 correctly spelled words**, and **45 total words written**.

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: $^\wedge$A$^\wedge$red$^\wedge$car

Capitalization

1. ONLY count capitalization as incorrect if capitalization is missing

   a. For the word "I"
   Example: ^Please^keep^in^mind^that^{X}i^{X}don't^like^you^.

   b. Proper names, like Jen, Florida, Santa Claus, Tooth Fairy
   Example: ^Each^year^he^starts^to^talk^about^{X}santa^{X}claus ^and^his^travels^across^the^world^.

   c. First word of sentence
   Example: ^{X}my^dog^ran^away^.

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 scoring like it's correct.

Example: ^My^Mom^likes^it^too^.

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).

Numerals

1.   Count numerals (i.e., 365) as ONE unit.

2.   Do not penalize for a lack of commas.

3.   Do not provide additional credit for commas.

    a.   Example: ^I^have^3,567^dogs^.

    b.   Example: ^I^have^3567^dogs^.

Punctuation

1. 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^birthday^is^January^12^,^1969^.

2. Do not provide additional credit for correctly used apostrophes (e.g., you'll, we'll, I'm, St. Patrick's).

    a.   Example: ^We've^been^eating^a^lot^of^food^.

    b.   Example: ^My^mom's^brother^went^to^school^.

3. Periods should be given credit when they are used correctly in the context of abbreviations.

    a.   Example: ^On^St^.^Patrick's^Day^I^celebrated^.

Grammar

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

   a. <u>Example</u>:   ^ <u>I</u> ^ <u>went</u> ^ <u>to</u> ^ <u>grandma</u> ^X <u>house</u>

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

   a. <u>Example</u>:   ^ <u>He</u> ^X <u>help</u> ^ <u>mom</u> ^ <u>in</u> ^ <u>the</u> ^ <u>kitchen</u>

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 and two incorrect writing sequences at this point.

   a. <u>Example</u>:   ^ <u>Murray</u> ^ <u>takes</u> ^ <u>the</u> ^ <u>train</u> ^ <u>to</u> ^ <u>school</u> ^X |
   ^X <u>Mom</u> ^ <u>rides</u> ^ <u>the</u> ^ <u>bus</u> ^ .

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. <u>Example</u>:   ^ <u>She</u> ^ <u>went</u> ^ <u>to</u> ^ <u>the</u> ^ <u>store</u> ^ <u>and</u> ^ <u>asked</u> ^   <u>for</u> ^
   <u>bread</u> ^X|^X <u>and</u> ^ <u>looked</u> ^ <u>at</u> ^ <u>books</u> ^ <u>and</u> ^   <u>went</u> ^ <u>home</u> ^ .

Spacing Issues

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

   a. <u>Example</u>:   ^<u>I</u> ^ <u>did</u>^ <u>my</u>^ <u>home</u> ^X <u>work</u>
2. If a student combines 2 words into 1 word ('alot' for 'a lot"; 'afew' for 'a few'; 'noone' for 'no one'), score this as incorrrect on both sides of the word, for example:

   a. <u>Example</u>:   ^<u>There</u>^ <u>were</u> ^X <u>alot</u> ^X <u>of</u> ^ <u>pencils</u>

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**, **ok, yeah**, **gotta, kinda**, are okay in dialogue only.

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

   a. <u>Example</u>: ^ <u>He</u> ^ <u>wore</u> ^X <u>like</u> ^X <u>a</u> ^ <u>t-shirt</u> ^ .

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.

<u>Example</u>:   ^<u>We</u> ^ <u>went</u> ^ <u>to</u> ^ <u>the</u> <u>sc</u>

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.

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

Ardoin, S. P., Eckert, T. L., & Cole, C. S. (2008). Promoting generalization of reading: A comparison of two fluency-based interventions for improving general education student's oral reading rate. *Journal of Behavioral Education*, *17*(3), 237-252.

Aud, S., Hussar, W., Johnson, F., Kena, G., Roth, E., Manning, E., Wang, X., and Zhang, J. (2012). *The Condition of Education 2012* (NCES 2012-045). U.S. Department of Education, National Center for Education Statistics. Washington, DC. Retrieved [date] from http://nces.ed.gov/pubsearch.

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

Bandura, A. (1969). *Principles of behavior modification*. Oxford England: Holt, Rinehart, & Winston.

Bates, T. C., Castles, A., Coltheart, M., Gillespie, N., Wright, M., & Martin, N. G. (2004). Behaviour genetic analyses of reading and spelling: A component processes approach. *Australian Journal Of Psychology*, *56*(2), 115-126.

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.

Blandford, B. J., & Lloyd, J. W. (1987). Effects of a self-instructional procedure on handwriting. *Journal Of Learning Disabilities*, *20*(6), 342-346.

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.

Buckley, J., & Cameron, L. D. (2011). Automaticity of exercise self-regulatory efficacy beliefs in adults with high and low experience in exercise self-regulation. *Journal Of Sport & Exercise Psychology*, *33*(3), 325-348.

Burns, M. K., & Symington, T. (2002). A meta-analysis of prereferral intervention teams: Student and systemic outcomes. *Journal Of School Psychology*, *40*(5), 437-447.

Bursuck, W. D., & Epstein, M. H. (1987). Current research topics in learning disabilities.

    Learning Disability Quarterly, *10*(1), 2-7.

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.

Calkins, L. (1994). *The art of teaching writing*. Portsmouth, NH: Heinemann Educational Books.

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, J., & Willis, J. (1978). Modifying components of creative behavior in the natural

    environment. *Behavior Modification*, *2*(4), 549-564.

Charlop, M. H., Schreibman, L., & Thibodeau, M. G. (1985). Increasing spontaneous verbal

    responding in autistic children using a time delay procedure. *Journal of Applied Behavior*

    Analysis, 18, 155-166.

Clariana, R. B., Wagner, D., & Murphy, L. (2000). Applying a connectionist description of

    feedback timing. *Educational Technology Research and Development*, *48*(3), 5-21.

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

    Ohio: Merrill.

Culham, R. (2005). *6+1 traits of writing for the primary grades*. Jefferson City, MO: Scholastic.

Cummings, K. D., Dewey, E. N., Latimer, R. J., & Good, R. (2011). Pathways to word reading

    and decoding: The roles of automaticity and accuracy. *School Psychology Review*, *40*(2),

    284-295.

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

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, *125*(6), 627-668.

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.

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.

Frye, E. M., & Gosky, R. (2012). Rapid word recognition as a measure of word-level automaticity and its relation to other measures of reading. *Reading Psychology*, *33*(4), 350-366.

Gould, Judith S. (1999). *Four Square Writing Method: A Unique Approach to Teaching Basic Writing Skills, Grades 4-6*. Crystal Springs Books.

Graham, S. (1983). The effect of self-instructional procedures on LD students' handwriting performance. *Learning Disability Quarterly*, *6*(2), 231-234.

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. 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., & 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.

Grskovic, J. A., & Belfiore, P. J. (1996). Improving the spelling performance of students with disabilities. *Journal Of Behavioral Education*, *6*(3), 343-354.

Gryiec, M., Grandy, S., & McLaughlin, T. F. (2004). The Effects of the Copy, Cover, and

    Compare Procedure in Spelling with an Elementary Student with Fetal Alcohol

    Syndrome. *Journal of Precision Teaching & Celeration*, *20*(1), 2-8.

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.

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

    classroom. Columbus, OH: Merrill.

Harper, G. F., Mallette, B., & Moore, J. (1991). Peer-mediated instruction: Teaching spelling to

    primary schoolchildren with mild disabilities. *Journal of Reading, Writing, & Learning*

    Disabilities International, *7*(2), 137-151.

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.

Harward, S. V., Allred, R. A., & Sudweeks, R. R. (1994). The effectiveness of our self-corrected

    spelling test methods. *Reading Psychology*, *15*(4), 245-271.

Hattie, J. A. (1999, June). Influences on student learning (Inaugural professorial address,

    University of Auckland, New Zealand). Retrieved from http://www.arts

    .auckland.ac.nz/staff/index.cfm?P=8650

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.

Helie, S., Roeder, J. L., & Ashby, F. (2010). Evidence for cortical automaticity in rule-based

    categorization. *The Journal of Neuroscience*, *30*(42), 14225-14234.

Hier, B. O., & Eckert, T. L. (2014). Evaluating elementary-aged students' abilities to generalize

    and maintain fluency gains of a performance feedback writing intervention. *School*

    Psychology Quarterly.

Hirsch, E. D. (2003). Reading comprehension requires knowledge—of words and the world:

    Scientific insights into the fourth-grade slump and the nations stagnant reading

    comprehension scores. *American Educator, 27*, 10–22, 28–29, 48.

Hooper, S. R., Knuth, S. B., Yerby, D., & Anderson, K. L. (2009). A review of science-

    supported writing instruction with implementation in mind. In S. Rosenfield, V.

Berninger (Eds.) , *Implementing evidence-based academic interventions in school settings* (pp. 49-83). New York, NY US: Oxford University Press.

Houghton Mifflin. (2013). Retrieved from English Grades K-5 website: http://www.eduplace.com/rdg/hme/k_5/index.html

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.

Hubbert, E. R., Weber, K. P., & McLaughlin, T. F. (2000). A comparison of Copy, Cover, and Compare and a traditional spelling intervention for an adolescent with a conduct disorder. *Child & Family Behavior Therapy*, *22*(3), 55-68.

Jackson, H. (1995, September). Performance feedback and generalization programming in writing instruction. *Dissertation Abstracts International Section A*, *56*.

Jaspers, K. E., Williams, R. L., Skinner, C. H., Cihak, D., McCallum, R., & Ciancio, D. J. (2012). How and to what extent do two cover, copy, and compare spelling interventions contribute to spelling, word recognition, and vocabulary development?. *Journal of Behavioral Education*, *21*(1), 80-98.

Johnson, K. R., & Layng, T. V. J. (1996). On terms and procedures: Fluency. *The Behavior Analyst, 19*, 281-288.

Jones, D., & Christensen, C. A. (1999). Relationship between automaticity in handwriting and students' ability to generate written text. *Journal Of Educational Psychology*, *91*(1), 44-49.

Karegianes, Myra L. Pascarella, Ernest T. Pflaum, Susanna W. ; The Journal of Educational Research, Vol 73(4), Mar-Apr, 1980. pp. 203-207.

Kazdin, A. E., & Bootzin, R. R. (1972). The token economy: An evaluative review. *Journal of Applied Behavior Analysis*, *5*, 343-372.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*(2), 254-284.

Koenig, E. A. (2014). Closing the feedback-standard gap with the use of a performance feedback and goal setting intervention to improve elementary students' writing fluency. *Dissertation Abstracts International Section A*, *74*.

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.

Kulhavy, R. W. (1977). Feedback in written instruction. *Review of Educational Research*, *47*(2), 211-232.

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.

Luu, P., Tucker, D. M., & Stripling, R. (2007). Neural mechanisms for learning actions in context. *Brain Research*, *1179*89-105.

Lysakowski, R. S., & Walberg, H. J. (1982). Instructional effects of cues, participation, and corrective feedback: A quantitative synthesis. *American Educational Research Journal*, *19*(4), 559-578.

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

Maheady, L., Harper, G. F., & Mallette, B. (1991). Peer-mediated instruction: A review of potential applications for special education. *Journal of Reading, Writing, & Learning Disabilities International*, *7*(2), 75-103.

Marholin, D., & Steinman, W. M. (1977). Stimulus control in the classroom as a function of the behavior reinforced. *Journal of Applied Behavior Analysis, 10,* 465-478.

Martens, B. K., & Collier, S. R. (2011). Developing fluent, efficient, and automatic repertoires of athletic performance. In J. K. Luiselli & D. R. Reed (Eds.), *Behavioral sport psychology: Evidence-based approaches to performance enhancement* (pp. 159-176). New York: Springer.

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.

Marzullo-Kerth, D., Reeve, S. A., Reeve, K. F., & Townsend, D. B. (2011). Using multiple-exemplar training to teach a generalized repertoire or sharing to children with autism. *Journal of Applied Behavior Analysis*, *44*(2), 279-294.

Mathson, D. V., Allington, R. L., & Solic, K. L. (2006). Hijacking fluency and instructionally informative assessments. In *Fluency instruction: Research-based best practices* (pp. 106-119). New York, NY: The Guilford Press.

Matson, J. L., Sevin, J. A., Box, M. L., & Francis, K. L. (1993). An evaluation of two methods for increasing self-initiated verbalizations in autistic children. *Journal of Applied Behavior Analysis*, *26*(3), 389-398.

McGuffin, M. E., Martz, S. A., & Heron, T. E. (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., & Espin, C. A. (2007). Technical features of curriculum-based measurement in writing: A literature review. *The Journal of Special Education*, *41*(2), 68-84.

McNeish, J., Heron, T. E., & Okyere, B. (1992). Effects of self-correction on the spelling performance of junior high school students with learning disabilities. *Journal of Behavioral Education*, *2*(1), 17-27.

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

Meindl, J. N., Ivy, J. W., Miller, N., Neef, N. A., & Williamson, R. L. (2013). An examination of stimulus control in fluency-based strategies: SAMFEDS and generalization. *Journal of Behavioral Education*, *22*(3), 229-252.

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.

Moser, L. A., Fishley, K. M., Konrad, M., & Hessler, T. (2012). Effects of the copy-cover-

    compare strategy on acquisition, maintenance, and generalization of spelling sight words

    for elementary students with attention deficit/hyperactivity disorder. *Child & Family*

    Behavior Therapy, *34*(2), 93-110.

Murphy, J. F., Hern, C. L., Williams, R. L., & McLaughlin, T. F. (1990). The effects of the copy,

    cover, compare approach in increasing spelling accuracy with learning disabled students.

    Contemporary Educational Psychology, *15*(4), 378-386.

National Center for Education Statistics (2012).The Nation's Report Card: Writing 2011 (NCES

    2012–470). Institute of Education Sciences, U.S. Department of Education, Washington,

    D.C.

National Writing Project. (2003). *Because writing matters: Improving student writing in our*

    schools. California: Jossey-Bass.

Nies, K. A., & Belfiore, P. J. (2006). Enhancing Spelling Performance in Students with Learning

    Disabilities. *Journal of Behavioral Education*, *15*(3), 163-170.

Pearson. (2009). *Wechsler Individual Achievement Test- Third Edition.* San Antonio, TX:

    PsychCorp.

Persicke, A., Tarbox, J., Ranick, J., & St. Clair, M. (2013). Teaching children with autism to

    detect and respond to sarcasm. *Research in Autism Spectrum Disorders*, *7*(1), 193-198.

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.

Peverly, S. T., Ramaswamy, V., Brown, C., Sumowski, J., Alidoost, M., & Garner, J. (2007).

    What predicts skill in lecture note taking?. *Journal of Educational Psychology*, *99*(1),

    167-180.

Pollard, J. S., Betz, A. M., & Higbee, T. S. (2012). Script fading to promote unscripted bids for

    joint attention in children with autism. *Journal of Applied Behavior Analysis*, *45*(2), 387-

    393.

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.

Raskind, W. H., Hsu, L., Berninger, V. W., Thomson, J. B., & Wijsman, E. M. (2000). Familial

    aggregation of dyslexia phenotypes. *Behavior Genetics*, *30*(5), 385-396.

Richards, T. L., Aylward, E. H., Berninger, V. W., Field, K. M., Grimme, A. C., Richards, A. L.,

    & Nagy, W. (2006). Individual fMRI activation in orthographic mapping and morpheme

    mapping after orthographic or morphological spelling treatment in child dyslexics.

    Journal of Neurolinguistics, *19*(1), 56-86.

Rivera, D., & Bryant, B. R. (1992). Mathematics instruction for students with special needs.

    Intervention in School and Clinic, 28(2*),* 71-86.

Saddler, B., Behforooz, B., & Asaro, K. (2008). The effects of sentence-combining instruction

    on the writing of fourth-grade students with writing difficulties. *The Journal Of Special

    Education*, *42*(2), 79-90.

Saddler, B., & Graham, S. (2005). The Effects of Peer-Assisted Sentence-Combining Instruction

    on the Writing Performance of More and Less Skilled Young Writers. *Journal of*

    Educational Psychology, *97*(1), 43-54.

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.

Samuels, S. J. (2006). Reading fluency: Its past, present, and future. In *Fluency instruction:*

    Research-based best practices (pp. 7-20). New York, NY: The Guilford Press.

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

    Inc.

Schulte-Körne, G. (2001). Genetics of reading and spelling disorder. *Journal Of Child*

    Psychology And Psychiatry, *42*(8), 985-997.

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.

Skinner, C. H., & Daly, E. (2010). Improving generalization of academic skills: Commentary on

    the special issue. *Journal of Behavioral Education*, *19*(1), 106-115.

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

    Fifth Edition. Boston: Pearson.

Spencer, V. G. (2006). Peer Tutoring and Students With Emotional or Behavioral Disorders: A

    Review of the Literature. *Behavioral Disorders*, *31*(2), 204-222.

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. F., Doud, C. L., Rowbury, T. G., & Baer, D. M. (1978). Peer facilitation of generalization in a preschool classroom. *Journal of Abnormal Child Psychology, 6,* 203-209.

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.

Tenenbaum, G., & Goldring, E. (1989). A meta-analysis of the effect of enhanced instruction: Cues, participation, reinforcement and feedback and correctives on motor skill learning. *Journal of Research & Development In Education*, *22*(3), 53-64.

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

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.

Vachon, F., & Jolicoeur, P. (2012). On the automaticity of semantic processing during task switching. *Journal Of Cognitive Neuroscience*, *24*(3), 611-626.

Ward, J., & Gow, L. (1982). Programming generalisation: A central problem area in educational psychology. *Educational Psychology*, *2*(3-4), 231-248.

Watt, J. M., & Topping, K. J. (1993). Cued spelling: A comparative study of parent and peer tutoring. *AEP (Association Of Educational Psychologists) Journal*, *9*(2), 95-103.

Wirtz, C. L., Gardner, R., Weber, K., & Bullara, D. (1996). Using self-correction to improve the spelling performance of low-achieving third graders. *Remedial And Special Education*, *17*(1), 48-58.

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

Zammuner, V. (1995). Individual and cooperative computer-writing and revising: Who gets the best results?. *Learning and Instruction*, *5*(2), 101-124.

Bridget Hier
Vita

Contact:
430 Huntington Hall
Syracuse, NY 13244

Education:

Syracuse University   2012
      Department of Psychology, M.S.

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