

ABSTRACT

BARNES, KRISTINA SAMANTHA. Monitoring and Measuring the Treatment Integrity of a School-Based Small Group Reading Intervention using Performance Feedback Procedures. (Under the direction of William P. Erchul and John C. Begeny.)

Four regular education teachers at an elementary school in a large district were trained to use the HELPS-SG reading intervention. Their implementation was observed over a period of ten to eleven-weeks. Implementation checklists were utilized while observing live HELPS-SG sessions in second and third grade regular education classrooms. Performance feedback procedures were then implemented in phases, occurring after the first three or five baseline sessions. Feedback was provided immediately following each session, and was both verbal and written.

Overall mean percent implementation adherence ranged from 86%-98% once feedback procedures began. There were moderate to significant increases in implementation adherence levels between the baseline and performance feedback phases. Additionally, interventionists required fewer sessions in order to implement the core steps of the intervention with high fidelity when compared to participants in the Begeny (2013) implementation study of the HELPS One-on-One program.

Overall findings suggest that the combined effect of verbal and written performance feedback procedures effectively increased interventionists' implementation of the HELPS-SG reading intervention.

Monitoring and Measuring the Treatment Integrity of a School-Based Small Group Reading
Intervention using Performance Feedback Procedures

by
Kristina Rodriguez Barnes

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Master of Science

Psychology

Raleigh, North Carolina

2014

APPROVED BY:

William P. Erchul
Committee Co-Chair

John C. Begeny
Committee Co-Chair

Scott A. Stage
Committee Member

BIOGRAPHY

Kristina S. Rodriguez-Barnes was born on August 11, 1986 in Manhattan, New York. In 1997, her family relocated to Pembroke Pines, Florida. In 2004, she graduated from Charles W. Flanagan High School. She completed her undergraduate studies at Florida State University in Tallahassee, Florida and received a Bachelor of Science degree in Psychology and Creative Writing in 2008. Kristina went on to work at a reading center for three years, where she had the opportunity to implement reading interventions at an elementary school on a Navajo reservation in New Mexico, before beginning her graduate career at North Carolina State University. She plans on pursuing her doctorate at North Carolina State University after obtaining her Master of Science degree in Psychology.

TABLE OF CONTENTS

LIST OF TABLES.....	vi
LIST OF FIGURES.....	vii
Introduction.....	1
Review of the Literature	3
Tiered Services and Evidence-Based Interventions in Schools	3
Treatment Integrity	5
Relevance of treatment integrity.....	5
Measuring treatment integrity.....	7
Increasing treatment integrity	8
Small Group Instruction.....	10
Helping Early Literacy with Practice Skills-Small Group Program.....	10
Conclusion	12
Statement of the Problem.....	13
Research Questions and hypothesis	14
Methods.....	15
Participants.....	15
Setting	16
Materials	16
HELPS-SG intervention and instructional materials	16
HELPS-SG implementation protocols.....	17
Monitoring HELPS-SG treatment integrity.....	17

Research Design	19
Experimental manipulations	20
Procedure	21
Interventionist training	21
Observer training	21
Protocol usage	22
Inter-rater agreement	22
HELPS-SG implementation	23
Performance feedback procedures	23
Maintenance procedure	24
Results.....	25
Inter-rater agreement	25
Descriptive Statistics.....	25
Research Questions 1 and 2 and Hypothesis 1	26
Fleur.....	27
Minerva.....	28
Molly	28
Ginny	29
Adherence to intervention steps	32
Research Question 3	38
Survey Responses	39
Discussion.....	41

Contextual factors	41
Discussion of individual cases	43
Fleur	43
Minerva	46
Molly.....	48
Ginny.....	49
Survey results.....	52
Limitations and Future Directions.....	54
Overall Conclusions and Implications.....	59
References.....	61
Appendices.....	68
Appendix A: Helping Early Literacy with Practice Skills- Small Group reading intervention (HELPS-SG) Instructional Materials.....	69
Appendix B: Helping with Early Literacy Practice Skills-Small Group reading intervention (HELPS-SG) Implementation Materials.....	73
Appendix C: Materials to Monitor Treatment Integrity using the Helping Early Literacy with Practice Skills-Small Group reading intervention (HELPS-SG).....	79
Appendix D: Teacher Consent Form.....	91
Appendix E: Teacher Questionnaire.....	95

LIST OF TABLES

Table 1	Treatment Integrity by Phase: Core Steps.....	25
Table 2	Treatment Integrity by Phase: Tips and Reminders.....	26
Table 3	Crosstabulation of Core Step Implementation.....	34
Table 4	Crosstabulation of Tips and Reminders Implementation	35
Table 5	Interventionist Survey Responses.....	40

LIST OF FIGURES

Figure 1 Display of Research Design 20

Figure 2 Percentage of Core Steps Implemented..... 30

Figure 3 Percentage of Tips and Reminders Implemented..... 31

Introduction

Systematically and methodically checking that all steps in an evidence-based intervention are completed – and completed correctly – increases the likelihood that an intervention is being carried out as intended, and therefore can be assumed to provide efficacious outcomes for individuals receiving the intervention. Monitoring and recording that an evidence-based intervention is being carried out accurately and in the order it was intended, is one aspect of treatment integrity. Theoretically, an intervention that is implemented with high integrity will be more effective than an intervention that is implemented with low integrity. When steps are missing or implemented incorrectly, an intervention cannot be assumed to be evidence-based, because missing or incorrectly implemented steps may change or nullify the effectiveness of the intervention (Perepletchikova, Treat, & Kazdin, 2007).

Additionally, the rise of Response to Intervention (RtI) has made the measurement of treatment integrity increasingly essential. A student's response to an evidence-based intervention within RtI determines the frequency and intensity of intervention and even their eligibility to undergo the process for special education services. If treatment integrity is not monitored and documented, the due process protections afforded to students become undermined (Sanetti & Kratochwill, 2009; Noell & Gansle, in press).

The purpose of the current study was to monitor and measure treatment integrity in a small-group format using written and verbal feedback procedures. Further exploration was warranted in this area not only because research on small group instruction is limited (Wasik, 2008), but also because research on the measurement of treatment integrity in small group

instruction is lacking. The current study used a reading fluency intervention that was adapted from a one-on-one format to a small-group format for use in a tier II setting.

Participating teachers in an elementary school received a two-hour workshop over two days, on the use of the intervention prior to implementing it. Trained research assistants then observed teachers and provided written and verbal feedback. A fading/maintenance process then occurred, where feedback gradually dissipated until teachers no longer received feedback. Teachers were also surveyed about the program as well as what was most and least helpful about the feedback procedures.

To support the current study, a review of current research in treatment integrity (TI) as well as the increasing use of evidence-based interventions (EBIs) in small group formats will be reviewed. First, a brief description of how EBIs are used within the schools will be given. Next, definitions of terms (e.g., treatment integrity, performance feedback) will be discussed, followed by the various methods used to measure treatment integrity. Then, an overview of available research on small group interventions and instruction will be provided, including gaps in the area of small group treatment integrity research. Finally, an in-depth discussion of the current study will be presented, including the methodology and results of the research questions and hypotheses.

Review of the Literature

Tiered Services and Evidence-Based Interventions in Schools

The Individuals with Disabilities Education Improvement Act (IDEIA, 2004) provides the option to use data obtained from EBIs to make eligibility decisions regarding learning disabilities (LDs) instead of an IQ/achievement discrepancy (Erchul & Martens, 2010; Jimerson, Burns, & VanDerHeyden, 2007; Kratochwill, Clements, & Kalymon, 2007). Response to Intervention (RtI) is a model that uses a problem-solving approach to make decisions about the effectiveness of an intervention structured within a multi-tiered system, using students' outcome data (Berkeley, Bender, Gregg Peaster, & Saunders, 2009; Kratochwill, et al., 2007; Reschly & Bergstrom, 2009).

RtI provides the opportunity for early identification of all children “at-risk” for academic failure or learning disabilities, and promotes the idea of prevention (Erchul & Martens, 2010). The RtI framework is comprised of a multi-tiered intervention service with instruction or intervention that is matched to the needs of students and becomes increasingly intensive as a child moves through them (Reschly & Bergstrom, 2009). Researchers are increasingly promoting the use of three tier models (e.g., Fuchs & Fuchs, 2006; Reschly & Bergstrom, 2009), and will therefore be the example provided for the purposes of this paper.

Within the three tiered model, tier 1, or universal prevention, comprises general classroom instruction and practices for all students and is intended to reduce new cases of problem behavior and/or academic failure (Erchul & Martens, 2010; Mitchell, Stormont, & Gage, 2011). This includes teaching of the core curriculum, routines for differentiating instruction, and positive behavior supports (Reschly & Bergstrom, 2009). Additionally, tier

1 classrooms include accommodations for all students, including those with disabilities. Screening measures and monitoring of students' performance in core instruction also occurs in this tier, and is meant to identify those children who are "at-risk" for academic failure (Fuchs & Fuchs, 2009).

Tier II, or secondary prevention, involves a standard form of small-group instruction that typically lasts 10-15 weeks (Fuchs & Fuchs, 2009). Instruction within this tier is moderately intensive and can focus on academic and/or behavioral interventions (National Center on Response to Intervention, 2010). Because tier II services include the identification of children who may have a disability, most students are expected to benefit from receiving evidence-based instruction when implemented with integrity. When data obtained from weekly progress monitoring suggest that a student is nonresponsive to instruction, the student is considered for more intensive instruction at tier III (Fuchs & Fuchs, 2009; Hughes & Dexter, 2011).

Tier III, or tertiary prevention, is the most intensive level of RTI. Instruction is often individualized and students meet one-on-one with teachers. Students are provided an intensified version of the EBI used in the previous tier or a new EBI. At this level, teachers must create end-of-year goals and increase progress monitoring as often as daily, in order to individualize instruction and adjust goals as needed (Erchul & Martens, 2010; Fuchs & Fuchs, 2009). Individuals who remain nonresponsive based on progress monitoring data at this level are then considered eligible for the special education referral process (Sanetti & Kratochwill, 2009).

Treatment Integrity

Treatment integrity is a complex and multidimensional construct. The term generally refers to four areas of an intervention: (a) content (i.e., the intervention steps delivered); (b) quality (i.e., how well the steps were delivered); (c) quantity (i.e., how much of the intervention was provided); and (d) process (i.e., how the intervention was delivered) (Sanetti & Kratochwill, 2009). Essentially, treatment integrity refers to the degree to which an intervention is implemented as intended within a conceptual model or manual (Sanetti & Kratochwill, 2009; Schulte, Easton, & Parker, 2009). Due to its multidimensional nature, a clear definition of treatment integrity is lacking, including a single definitive term. For example, treatment integrity may be referred to as treatment fidelity, intervention integrity, and implementation fidelity. One also must consider how to parse out and identify specific constructs within treatment integrity (e.g., adherence, quality, process) when discussing it. For the purposes of this paper, *treatment integrity* will be used to describe two constructs, specifically: adherence and quality, as these aspects remain the focal point for school-based practice (Schulte et al., 2009).

Relevance of treatment integrity. Treatment integrity has become relevant to practitioners and researchers in education due to a variety of factors, including federal legislation such as No Child Left Behind (2001) and the Individuals with Disabilities Education Improvement Act (2004). These laws hold educators accountable for best practices and the proper use of EBIs before a child may be considered eligible to undergo the special education referral process (Sanetti & Kratochwill, 2009). Because of this, a growing number of professional organizations, such as the National Association of School

Psychologists, have become increasingly interested in the measurement and maintenance of treatment integrity in research and practice (Sanetti & Kratochwill, 2009).

To review the opinions of practitioners in the field regarding treatment integrity, Cochrane and Laux (2008) conducted a survey of 806 Nationally Certified School Psychologists. They found that participants believed the measurement of treatment integrity was a key factor to include when evaluating interventions. The investigators further documented that of the total sample, 91 (11.3%) indicated they always measured TI in one-on-one consultation with teachers and/or parents, 335 (41.6%) said they measured TI “sometimes,” and 270 (33.5%) responded with “no-never.” Those who indicated “yes” or “sometimes” generally relied on indirect methods such as teacher self-report or interview 44.8% of the time and often after the intervention had been implemented. Additionally, participants were surveyed about problem-solving team (PST) records. Specifically, they were asked if PST records were reviewed, and, if so, the percentage of the time a statement or numerical index was available, indicating evidence of TI measurement. Participants reported that 67.3% of the time, team records provided no documentation of TI measurement.

Cochrane and Laux’s (2008) survey concluded by asking participants their opinions regarding the importance of TI in school-based interventions (e.g., why or why not they believed it was important). Participants identified four main factors as to why treatment integrity data were not collected: (a) lack of time and resources; (b) lack of teacher, staff, and administrator knowledge of its importance; (c) lack of administrator and/or system support for the collection of treatment integrity data; and (d) lack of acceptance by teachers for collection of treatment integrity data. These four areas of difficulty have been echoed across

studies attempting to measure treatment integrity, with indirect and direct methods of measurement occurring most often. These methods are described next.

Measuring treatment integrity. Indirect methods of measurement involve self-reports, rating scales, and interviews (Cochrane & Laux, 2008). Although these methods are less intrusive, they are more subjective because they are all forms of self-report. Additional indirect methods include the use of permanent products and manualized interventions. Permanent products are tangible results generated from each step or component of an intervention (e.g., a sticker chart) whereas manualized interventions have step-by-step guidelines or instructions for implementing the intervention. However, each measure has its drawbacks. Permanent products require that each step in a given intervention must produce a tangible product, whereas treatment integrity is often assumed and not systematically measured in manualized interventions (Cochrane & Laux, 2008; Lane, Bocian, MacMillan, & Gresham, 2004).

Direct methods of measurement involve observing the person implementing the intervention and recording whether each step of the implementation is present or absent (e.g., Cochrane & Laux, 2008; Gresham, McMillian, Beebe-Frankenberger, & Bocian, 2000; Lane et al., 2004). Direct methods of observation are the most objective and can provide the most accurate assessment of treatment integrity (Cochrane & Laux, 2008). However, this method might cause reactivity on the part of the person being observed, and could be inappropriate or impossible to use for all academic and behavioral interventions (Cochrane & Laux, 2008; Noell, 2008; Wilkinson, 2006).

Direct methods can also be time consuming. Lane et al. (2004) reviewed four key steps that need to be followed in order to conduct direct observations. First, if there is no treatment integrity protocol, one must be created. A treatment integrity protocol is a detailed list or task analysis of intervention components (Lane et al., 2004). Next, operational definitions of the intervention components must be created. For example, “student response” could be defined as a student giving either a verbal response (e.g., no) or physical gesture (e.g., thumbs up) to indicate agreement or disagreement with a given statement. Then, presence or absence of each step or component is documented. Finally, treatment integrity is assessed and reported in the form of percentages. For example, if 8 of 12 steps in an intervention were completed, then the treatment integrity for that session would be 67%. In other words, the intervention was carried out with 67% integrity or was only carried out 67% accurately and as prescribed.

Increasing treatment integrity. Performance feedback is a popular method used to increase treatment integrity of both behavioral and academic classroom-based interventions (DiGennaro, Martens, & McIntyre, 2005; Solomon, Klein, & Politylo, 2012). This method typically involves a person who is trained in the intervention components yet does not implement the intervention. The person observing, often referred to as a consultant, then observes the intervention implementer and provides corrective feedback once the session is over (Lane et al., 2004). Additionally, performance feedback can be modified to suit the unique needs of the consultant-consultee dyad. For example, the immediacy in which a consultant provides feedback can be tailored to the need of the consultee to be both

immediate for core steps and delayed for overall performance feedback (Solomon et al., 2012).

Solomon et al. (2012) conducted a meta-analysis and found that performance feedback yields significant effects regardless of setting, dependent variable, delay of feedback, or type of intervention. However, they also found that the effect of performance feedback is greater for increasing the treatment integrity of academic interventions than behavioral interventions. They suggested that additional research is needed to consider manipulating the delay of feedback and analyzing different components of feedback, such as using graphs or modeling the intervention in order to determine which types of performance feedback are most effective.

Although the Solomon et al. (2012) study revealed important information about performance feedback, this meta-analysis focused primarily on dyadic interventions and whether interventions were behavioral or academic. Little research has been conducted in terms of examining treatment integrity methods, such as performance feedback in small group interventions. Exploring this gap is crucial in determining if treatment integrity methods that work in dyadic interventions are just as effective for small group interventions.

Small Group Instruction

Small group instruction has been found to be effective as a means to support “at-risk” populations in reading and in early childhood classrooms (Foorman & Torgesen, 2001; Wasik, 2008). This format for instruction is increasingly being suggested for use within RtI, particularly tier II services (Fuchs & Fuchs, 2009). When implemented correctly, it might be assumed that interventions designed for small group instruction can provide schools the

opportunity to maximize their resources in order to deliver treatment to a larger group of students more effectively and efficiently (Foorman & Torgesen, 2001; Wasik, 2008). It is therefore important to be sure the integrity of an intervention is systematically measured to ensure that students receive the best care possible, as well as to ensure that other proven methods of improving TI (e.g., performance feedback) work in the context of a small group intervention. An emerging body of literature is now focusing on the importance of measuring and enhancing TI within all tiers of RtI. Various methodologies have been suggested and implemented to promote monitoring of TI within each tier, including self-reports and various forms of performance feedback (e.g., verbal, graphical, TI protocols) (e.g., Barnett, Hawkins, & Lentz, 2011; Bradshaw, Debnam, Koth, & Leaf, 2008; Mayer, 2012).

Helping Early Literacy with Practice Skills-Small Group Program

The Helping Early Literacy with Practice Skills-Small Group program (HELPS-SG) is an example of an academic, small group intervention that can be implemented in schools as an oral reading fluency intervention. This intervention includes a standardized flow chart of procedural steps and scripted instructions. There is no treatment integrity protocol specific to this intervention yet; however, due to the similarity between the HELPS-SG program and the HELPS one-on-one program (Begeny 2009), a treatment integrity protocol can easily be adapted from the HELPS one-on-one program to the HELPS-SG program.

In a study conducted by Begeny, Easton, Upright, Tunstall, and Ehrenbock (2013), the reliability and user-feasibility of the HELPS one-on-one treatment integrity protocol (also referred to as implementation integrity monitoring system) were reviewed. All participants

were required to independently follow the self-training module described in the HELPS Program Teacher's Manual (Begeny, 2009, pp. 19-36). Participants were required to be HELPS interventionists prior to being eligible as observers. There were two sets of observers, referred to as observer 1 and observer 2. The criteria for being observer 1 included the ability to implement the core HELPS procedures with 100% integrity at least three times within five consecutive sessions and demonstrate 90% or better integrity using the Tips and Reminders protocol at least three times within five consecutive sessions. In order to become observer 2, one first had to be an observer 1 and achieved 100% accuracy on the Steps for Observing form (which includes accurate completion of two additional TI protocols) at least three out of five consecutive observations, with an overall accuracy always above 80% during the five consecutive sessions. Observer 1's did not receive performance feedback, but accuracy was calculated by using inter-observer agreement from observer 2's.

Begeny et al. found that with a combination of their implementation materials and feedback procedures, new interventionists were able to reach the stringent implementation criteria (i.e., 100% implementation of core procedures three times within five consecutive sessions and 90% implementation of Tips and Reminders three times within five consecutive sessions). Because these stringent criteria were met with novice implementers successfully, it is believed that these criteria are appropriate for experienced intervention teachers.

Although the Begeny et al. (2013) findings are promising, it is important to note that no measurement of TI regarding the HELPS-SG program has yet been conducted. As small group instruction in tier II services increases, it is important to evaluate the TI of any interventions before they are to be considered evidence-based and used in schools.

Conclusion

In summary, response to intervention is a multi-tiered framework that employs a problem-solving approach in order to address varying levels of student need based on universal screening and targets those students with gaps between expected and actual performance (Reschly & Bergstrom, 2009). Additionally, RtI is an alternative framework for making eligibility decisions for special education (Kratochwill, et al., 2007). The structure of RtI promotes varying degrees of intensive services, where students have access to increased instruction based on their level of need. Within each tier, a problem-solving approach is used to determine exactly what type of instruction a student needs as well as when it is time for a student to undergo the special education eligibility process. Small group interventions are increasingly used in schools as part of these tiered services, particularly within the second tier, and are seen as an integral solution in maximizing already limited school resources (Wasik, 2008). As the trend to use small group EBIs continues, it is important to measure and monitor the treatment integrity of these interventions because moving a student throughout tiers of RtI without documenting and measuring the treatment integrity of a given intervention denies students of their due process rights (Noell & Gansle, in press). If a student does not respond to an intervention, it should be due to an increased need of services, not due to the improper implementation of an intervention.

Statement of the Problem

The following points summarize the areas of concern in school-based intervention implementation of adherence monitoring and support methods:

- The rise of RTI and tiered services in schools requires the use of EBIs in order to support children labeled as “at-risk” for behavior and/or academic disabilities. Measurement and maintenance of treatment integrity (TI) is central to sustain the effectiveness of evidence-based interventions.
- TI is essential to achieve and document to make valid conclusions regarding treatment outcome. However, there is little research about how TI is measured in tier II small group interventions.
- Performance feedback – a proven method of increasing treatment integrity – is a crucial factor that affects treatment integrity and maintenance. However, little is known about how performance feedback procedures affect treatment integrity of tier II small group interventions.
- HELPS-SG is an example of an evidence-based small group reading intervention that can be used for tier II instructional support. Although there is documentation of high levels of TI when using specific integrity protocols for the one-on-one program, there is no documentation of TI for the small group program, nor is there a TI protocol available.

Research Questions

The research questions (RQs) and hypothesis (H1) of the current research study were as follows:

1. To what extent does the combined effect of verbal and written feedback affect treatment integrity of a tier II small group intervention, such as HELPS-SG, when steps and scripted instructions are provided as part of the intervention materials?
 1. H1: There will be a positive trend with treatment integrity increasing once feedback procedures begin.
2. To what extent are feedback procedures necessary to maintain a high level of treatment integrity (i.e., 100% implementation of core steps, 90% implementation of supplemental steps) over time in a small group intervention?
3. Are there differences in the amount of time it takes for participants to reach high levels of treatment integrity compared to the study conducted by Begeny et al. (2013)?

Methods

Participants

Participants included eight elementary school teachers drawn from a central North Carolina elementary school in a large district. Two teachers removed themselves before the training session occurred (25%). Once the study began, two teachers removed themselves from proceeding (25% attrition) for various reasons. The first participant was ruled a poor fit after the first workshop due to her infrequent focus instruction time, and the second participant removed herself two sessions into the study, citing insufficient time to participate. A final total of four participants remained in the study (50%).

Grades ranged between second and third and a total of 12 students received the reading fluency intervention. Teachers were identified for the study by the principal of the school. Students were identified based on universal screening data (e.g., Dynamic Indicators of Basic Literacy Skills (DiBELS); Kaminski & Good, 2004) and teacher selection of students who would most benefit from intervention. DiBELS scores ranged from 57 words per minute to 102 words per minute. Students were grouped together based on DiBELS (Kaminski & Good, 2004) data and the HELPS placement test. Demographic information regarding teachers is as follows: a) teacher one (Fleur) was a second grade regular education teacher, and has been teaching for 13 years (2 years in second grade); b) teacher two (Minerva) was a second grade regular education teacher, and has been teaching for 13 years (11 years in second grade); c) teacher three (Molly) was a third grade regular education teacher, and has been teaching for 15 years (2 years in third grade) and; d) teacher four (Ginny) was a third grade regular education teacher, and has been teaching for 8 years (first

year in third grade). Observers were trained undergraduate and graduate research assistants from the North Carolina State University, including the principal investigator.

Because all participants implemented the intervention, they will henceforth be referred to as interventionist(s). All observers were trained in the HELPS-SG intervention prior to the start of data collection and were undergraduate research assistants, graduate students, or the principal investigator. Observers in this study were responsible for: (a) watching and recording interventionists implement the HELPS-SG reading intervention; (b) recording all implementation data using the appropriate HELPS-SG TI protocols (discussed in further detail below); and (c) providing verbal and written feedback to interventionists when appropriate (discussed in further detail below).

Setting

All treatment integrity monitoring activities and observations associated with this study occurred in the classroom of each of the participating interventionists. Students received intervention during focus instruction time.

Materials

HELPS-SG intervention and instructional materials. The HELPS-SG instructional procedures include eight evidence-based strategies (e.g., repeated-reading, motivational strategies) shown in previous research to improve students' reading fluency (Begeny et al; 2013). The HELPS-SG reading intervention was selected for the purposes of this study because it offers a set of materials that guide an interventionist through implementation. The HELPS-SG reading intervention includes: (a) specific implementation protocol for an interventionist to follow while implementing HELPS-SG procedures (See

Appendix A); (b) an individual Progress Tracking Form for the interventionist to document implementation of HELPS-SG for individual students and facilitate communication between different interventionists who may work with the same student across different days; (c) an individual Student Graph to document implementation, employ goal setting, and performance feedback procedures; (d) a Group Progress tracking form for the interventionist to document implementation of HELPS-SG; and (e) a Group Star Chart as the program's motivational/reward system.

HELPS-SG implementation protocols. There are four main HELPS-SG implementation protocols that were used to train the interventionist to implement the program and used to develop the treatment integrity monitoring materials. The protocols included: (a) a detailed Implementation Protocol that describes the 13 core implementation steps and their proper sequence (see Appendix A); (b) a Scripted Directions protocol that specifies the verbal directions that should be used at each step of the intervention; (c) the Implementation Flow Chart that summarizes the core steps of the program in their correct sequence; and (d) the Tips and Reminders for Implementation protocol (see Appendix C), which lists different minor steps for implementation and is designed to improve the quality of the implementation (i.e., factors that extend beyond the interventionist following the 12 core steps). All implementation protocols (except the Tips and Reminders for Implementation protocol) can be found in Appendix B.

Monitoring HELPS-SG treatment integrity. There are four documents that were developed from the HELPS one-on-one intervention for this study. These documents were intended to facilitate an experienced HELPS one-on-one observer to successfully monitor

and record an interventionist's implementation integrity. Each of these forms will be described next and can be found in Appendix C.

The *Steps for Observing Teachers During their Implementation of the HELPS-Small Group Program* (or the *Steps for Observing form*) was adapted from the HELPS one-on-one program and includes specific steps that should be used when observing an interventionist, including how to provide feedback to the interventionist about his or her implementation integrity and provided a reminder to the observer to record the interventionist's implementation integrity on the *Observation Summary Form* and the *HELPS Observation Checklist* (described later in this section). The steps of the *Steps for Observing form* for the HELPS one-on-one program were based on "research describing effective practices for training, observing, and providing effective practices for training, observing, and providing effective feedback to educators in the context of intervention implementation" (Begeny et al., 2009, p. 10). Next, the *Observation Summary Form*, the *HELPS-SG Observation Checklist*, and the *HELPS Group-Based Program: Project Procedural Integrity* protocols will be discussed.

The *Observation Summary Form* (OSF) required that observers document the following information: (a) observed interventionists' names; (b) observers' names; (c) date(s) of observations; (d) observation session number(s) (based on the number of times interventionists have been observed); (e) group information (e.g., student names, HELPS-SG session number); (f) ratings (on a 1-5 Likert scale) of the interventionists' organization and enthusiasm with students; (g) percentages of core HELPS-SG steps and Tips and Reminders implemented correctly; (h) lists of core steps not implemented correctly (when appropriate);

(i) relevant notes (e.g., Tips and Reminders not implemented correctly); (j) questions or concerns raised by interventionists during post-session discussions; (k) steps for interventionists to focus on implementing correctly during subsequent HELPS-SG sessions; (l) whether all interventionists' questions were addressed during post-session feedback; and (m) whether all the missed or incorrectly implemented core steps were reviewed during post-session feedback discussions (Begeny et al., 2009, p. 12).

The *HELPS-SG Observation Checklist* (Observation Checklist) was a form that specified each of the core implementation steps of the HELPS-SG program as well as each of the items from the Tips and Reminders Checklist. The *HELPS Group-Based Program: Project Procedural Integrity Form* (also referred to as the Procedural Integrity Form) was a document that was filled out by the observer with the following information: (a) the name of the observer; (b) date of observation; (c) the name of the interventionist; (d) the number of core steps completed correctly (e.g., 8 of 12, 12 of 12); and (e) quick notes about the observation or steps not completed. Unlike the Observation Checklist, the Procedural Integrity Form focuses on only one interventionist and the 13 core steps of the HELPS-SG program and serves as a shorthand version of the Observation Checklist and OSF for note-taking purposes.

Research Design

This study used a multiple baseline across subjects design. Specifically, this design included three phases, which will result in an A-BC-D design. The dependent variable is TI of the interventionists, with the independent variable being verbal and written performance feedback. The first phase (A) consisted of gathering baseline data. The second phase (BC)

consisted of implementing both verbal and written performance feedback to interventionists until they reach the specified integrity criteria (i.e., 100% of core steps three times within five consecutive sessions, 90% of Tips and reminders at three times within five consecutive sessions).

Feedback procedures occurred in the interventionists' respective classrooms after every observed HELPS-SG session three times a week. The final phase (D) began once the treatment integrity criteria had been met. During this time performance feedback was given once a week, as needed (i.e., falling below TI criteria) (DiGennaro et al., 2007). This phase was intended to be a maintenance phase and concluded the study. Additionally, interventionists were surveyed at the conclusion of the study. The survey included items such as: (a) which feedback procedures were most effective; (b) which feedback procedures were least effective; (c) which steps of the HELPS-SG program were easiest to implement; and (d) which steps of the HELPS-SG program were most difficult to implement. 1 provides a graphical display of the research design.

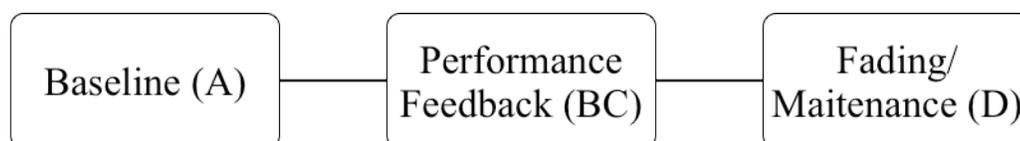


Figure 1. Display of Research Design

Experimental manipulations. Because the focus of this study was on the effect of verbal and written feedback procedures on a small group intervention, the investigator

staggered the times in which participating interventions begin receiving feedback once baseline had been established. Baseline was established using visual analysis. Once baseline was established, interventionists were randomly assigned into different feedback conditions that determined when an interventionist began receiving feedback (e.g., immediately after baseline, after 4 sessions, or after 6 sessions).

Procedures

Interventionist training. Participating interventionists were given a consent form to review (see Appendix D). Upon signing the consent form, interventionists received an initial 45-minute overview of the intervention, and a second one and a half hour workshop per grade on the use and implementation of the HELPS-SG program. At the initial workshop, interventionists received all materials and forms necessary to implement the HELPS-SG program. Main topics discussed included usage of the procedural flow chart, forms (e.g., Start Chart, Group Tracking Form), scripted instructions, and a sample video of a graduate research assistant implementing the program with volunteers. The second workshop included the opportunity for interventionists to practice using HELPS-SG forms by role-playing with other interventionists and/or the trainer (i.e., the principal investigator). At the conclusion of the training, interventionists kept all practice materials as a reference (except the video), as well as the necessary binders containing the required student and teacher passages, and all additional implementation materials (e.g., group tracking form, group star chart).

Observer training. Observers had been previously trained in the HELPS one-on-one program and certified under rigorous criteria (i.e., 100% implementation of core steps in

three of five consecutive sessions and 90% and above implementation of tips and reminders in three of five consecutive sessions). Observers received one week of training in the HELPS-SG program to ensure thorough understanding of proper intervention implementation and comfort with feedback procedures, including use of all forms and protocols. During training, observers had the opportunity to practice using the HELPS-SG intervention via role-play and practiced using the implementation and observation protocols as well as giving performance feedback. Live HELPS-SG sessions were performed by the principal investigator and followed a role-playing model.

Protocol usage. The following forms were used by observers: (a) *The Steps for Observing Teachers During their Implementation of the HELPS-Small Group Program* (or the *Steps for Observing Form*); (b) the *Observation Summary Form*; (c) the *HELPS Observation Checklist*; and (d) the *HELPS Group-Based Program: Project Procedural Integrity Form*.

Inter-rater agreement. As previously stated, during observer training, observers had the opportunity to watch the HELPS-SG program be implemented while scoring TI using the required forms. Additionally, they practiced providing verbal feedback after watching these sessions. These practice sessions were also used to test inter-rater reliability. Each undergraduate assistant had to meet 100% agreement during practice sessions with the principal investigator. All six raters reached 100% agreement on three separate practice sessions. Once interventionists began implementing HELPS-SG, observers recorded the implementation of the intervention and began feedback procedures during the appropriate phases. All sessions were recorded via audio recorders. The percentage of recorded

implementation sessions for agreement checks is greater in this study than that used by previous researchers (e.g., Burns, et al. 2008; Mayer, 2012). The principal investigator reviewed and scored TI in order to obtain inter-rater reliability using the required TI protocols. Inter-rater agreement was measured by counting the number of agreements divided by the total number of steps and multiplied by 100. It was expected that 90% inter-rater agreement would be reached. 3 observers fell below 90% agreement and required one to two additional training sessions.

HELPS-SG implementation. Interventionists implemented the HELPS-SG program three times a week as recommended per the HELPS-SG author, for a total of six weeks. HELPS-SG occurred during interventionists' regularly scheduled language arts focused instruction time. Students in their groups were identified through start of term universal screening measures such as the Dynamic Indicators of Basic Early Literacy (DiBELS; Kaminski & Good, 2004) and teacher identification of students whom they believed would most benefit from additional instruction.

Performance feedback procedures. Both verbal and written performance feedback were provided to each interventionist once baseline has been established. Interventionists were randomly assigned to one of two groups, with each group receiving feedback at different points. Specifically, teachers were grouped to receive feedback after three or five sessions, based on visual analysis. The sessions and their amounts were chosen due to time constraints of the participants. A copy of the *HELPS Observation Checklist* was provided to interventionists after each feedback session to fulfill the written feedback component of the proposed study. Verbal feedback was given using the *Steps for Observing Teachers During*

their Implementation of the HELPS-Small Group Program (or the Steps for Observing Form) and the *Observation Summary Form*. All feedback was given immediately following all HELPS-SG sessions. No feedback or modeling occurred while HELPS-SG was being implemented.

Maintenance procedure. Once an interventionist reached 100% implementation of core procedures in three of five consecutive sessions and 90% or above in all Tips and Reminders, a fading/maintenance process began. Maintenance procedures included reducing feedback procedures from three times a week to once a week for two to three weeks. Feedback was provided as needed (i.e., falling below the aforementioned TI standards). Additionally, a survey regarding program usage and helpfulness of performance feedback procedures was provided. For example, interventionists were asked which steps were easiest to master and which were more difficult, or what type of performance feedback they prefer (e.g., verbal, written, or both).

Results

Inter-rater Agreement

Agreement was calculated by the number of agreements divided by the total amount of observable steps for all sessions, multiplied by 100. Agreement checks between all research assistants were conducted for all observable intervention sessions, which resulted in a range of 69%-100% agreement on Core Steps and 70%-100% agreement on Tips and Reminders, with an average agreement of 93.5% on Core Steps and 96.3% on Tips and Reminders.

Descriptive Statistics

Means, standard deviations, and medians of all steps implemented (i.e., Core Steps and Tips and Reminders) were calculated. Tables 1 and 2 present descriptive data for the implementation of both Core Steps and Tips and Reminders per interventionist and include all means, medians, and standard deviations for both the Core Steps and the Tips and Reminders quality steps.

Table 1 *Treatment Integrity by Phase: Core Steps*

Participants	Implementation at baseline			Implementation after performance feedback		
	<i>M</i>	<i>SD</i>	<i>Mdn</i>	<i>M</i>	<i>SD</i>	<i>Mdn</i>
Fleur	56.41	11.75	54.00	86.33	17.53	100
Minerva	56.41	8.89	61.54	89.74	10.88	76.92
Molly	83.08	12.64	100.00	98.46	3.44	100
Ginny	55.77	9.68	57.69	95.39	6.88	100

Table 2 *Treatment Integrity by Phase: Tips and Reminders*

	Implementation at baseline			Implementation after performance feedback		
	<i>M</i>	<i>SD</i>	<i>Mdn</i>	<i>M</i>	<i>SD</i>	<i>Mdn</i>
Participants						
Fleur	50.47	7.55	44	77.14	12.56	84.21
Minerva	58.5	20.39	73.53	90.59	6.87	91.67
Molly	56.82	14.99	64.71	89.58	7.73	94.12
Ginny	48.3	7.27	49.82	77.88	6.42	72.73

Research Questions One and Two and Hypothesis One

In order to determine to what extent the combined effect of verbal and written feedback affect treatment integrity of a tier II small group intervention, interventionists' TI was analyzed via three methods: (a) visual analysis, which has been deemed acceptable for evaluating single-subject data (Kaufman, Coddling, Markus, Shick-Tryon, & Nagler-Kyse, 2013; Kazdin, 2001), (b) percentage of non-overlapping data, and (c) a cross-tabulation of all steps analyzed by the Pearson chi-square test.

Overall, once baseline was established and PFB procedures began, all interventionists improved their implementation of the HELPS-SG intervention (see Figure 2). At baseline, interventionists' mean implementation of Core Steps ranged from 55.77 - 83.07% (*Mdn* = 69.75%) in their implementation of the intervention. However, as expected, interventionists' treatment integrity increased overall once performance feedback procedures began. After the first day of PFB measures, interventionists' mean implementation ranged from 46%-100% (*Mdn* = 84.5%) . By the end of the study, interventionists' mean implementation ranged from 86.33% to 98.46% (*Mdn* = 88.46%). There were some reported drops in TI during

performance feedback procedures, which will be further reviewed in the Discussion. Next, individual interventionists' data are presented.

Fleur. Fleur's baseline implementation of the HELPS-SG intervention Core Steps ranged from 46%-70% adherence ($M = 56.41\%$; $Mdn = 54\%$). Once performance feedback procedures began, implementation of Core Steps ranged from 46%-100% adherence ($M = 86.33\%$, $Mdn = 100\%$). It should be noted that there were two sessions where implementation fell below 100% during feedback procedures. This will be addressed in the Discussion.

Baseline implementation of the HELPS-SG Tips and Reminders ranged from 44-59% adherence ($M = 50.3$; $Mdn = 44\%$). Implementation of the Tips and Reminders ranged from 55.88%-88.24% ($M = 77.14\%$; $Mdn = 84.21\%$) once performance feedback procedures began. It should be noted that there were two sessions where implementation dropped compared to other sessions and this will be reviewed in the Discussion.

Finally, in order to determine to what extent the combined verbal and written feedback affected treatment integrity, the Percentage of Non-Overlapping Data (PND) statistic was calculated. Fleur's PND of Core Steps was 89% and is considered moderately effective based on Scruggs and Mastropieri's rubric for interpreting effect sizes (Lentz, 2012; Scruggs & Mastropieri, 1998). The PND was also calculated for Fleur's implementation of the HELPS-SG Tips and Reminders. Her PND statistic of 78% indicates a moderate effect. Overall, the combined effects of verbal and written feedback were moderately effective in improving Fleur's implementation of the HELPS-SG reading intervention.

Minerva. Minerva's baseline implementation of the HELPS-SG intervention Core Steps ranged from 61.54%-61.54% ($M = 56.41\%$; $Mdn = 61.54\%$). Once performance feedback procedures began, implementation of Core Steps ranged from 69.23%-100% adherence ($M = 89.74$; $Mdn = 76.92\%$). Baseline implementation of the HELPS-SG Tips and Reminders ranged from 35.29%-73.53% ($M = 58.50\%$; $Mdn = 73.53\%$). Implementation of the Tips and Reminders ranged from 73.53%-97.06% ($M = 90.59\%$; $Mdn = 91.67\%$) once the performance feedback phase began.

Minerva's PND of Core Steps was 100% and is considered very effective while the PND for the HELPS-SG Tips and Reminders quality steps of 89% indicates a moderate effect. Overall, the combined effects of verbal and written feedback ranged from moderate to very effective for Minerva.

Molly. Molly's baseline implementation of the HELPS-SG intervention Core Steps ranged from 69.23%-100% ($M = 83.08\%$; $Mdn = 100\%$). Once performance feedback procedures began, implementation of Core Steps ranged from 92.31%-100% ($M = 98.46\%$; $Mdn = 100\%$). There were two sessions where implementation of Core Steps fell below 100% during the feedback phase. Baseline implementation of the HELPS-SG Tips and Reminders quality steps ranged from 35.29%-75.76% ($M = 56.82\%$; $Mdn = 64.71\%$). Implementation of the Tips and Reminders ranged from 76.47%-94.44% ($M = 89.58\%$; $Mdn = 94.12\%$) once the performance feedback phase began.

Molly's PND of Core Steps was 0% and obviously was not effective. However, one drawback of the PND statistic as used here is its reliance on only one data point in the baseline phase and is thus vulnerable to outliers. Molly did have an outlier session that

constituted the baseline point used to calculate the PND based on Scruggs and Mastropieri's (1998) description. Therefore, the effect of the feedback intervention cannot reliably be calculated. The PND was also calculated for Molly's implementation of the HELPS-SG Tips and Reminders. The PND statistic of 83% indicates a moderate effect. Overall, although effects could not be calculated for the Core Steps, the combined effects of verbal and written feedback were found to be moderately effective for the Tips and Reminders.

Ginny. Ginny's baseline implementation of the HELPS-SG intervention Core Steps ranged from 46.15%-69.23% ($M = 55.77\%$; $Mdn = 57.69\%$). It should be noted that Ginny had one missing baseline session thus her baseline consisted of only four data points instead of five. Once performance feedback procedures began, implementation of Core Steps ranged from 84.62%-100% ($M = 95.34\%$; $Mdn = 100\%$). Additionally, there was one session where the implementation of the Core Steps fell below 100% during the feedback phase. Baseline implementation of the HELPS-SG Tips and Reminders quality steps ranged from 40.63%-55.88% ($M = 48.30\%$; $Mdn = 49.82\%$). Implementation of the Tips and Reminders ranged from 72.73%-88.57% ($M = 77.88\%$; $Mdn = 72.73\%$) once the performance feedback phase began.

Ginny's PND of Core Steps was 83% and is considered moderately effective, while her PND statistic of 100% on Tips and Reminders indicates that performance feedback was very effective. Overall, the combined effects of verbal and written feedback ranged from moderately effective to highly effective for her.

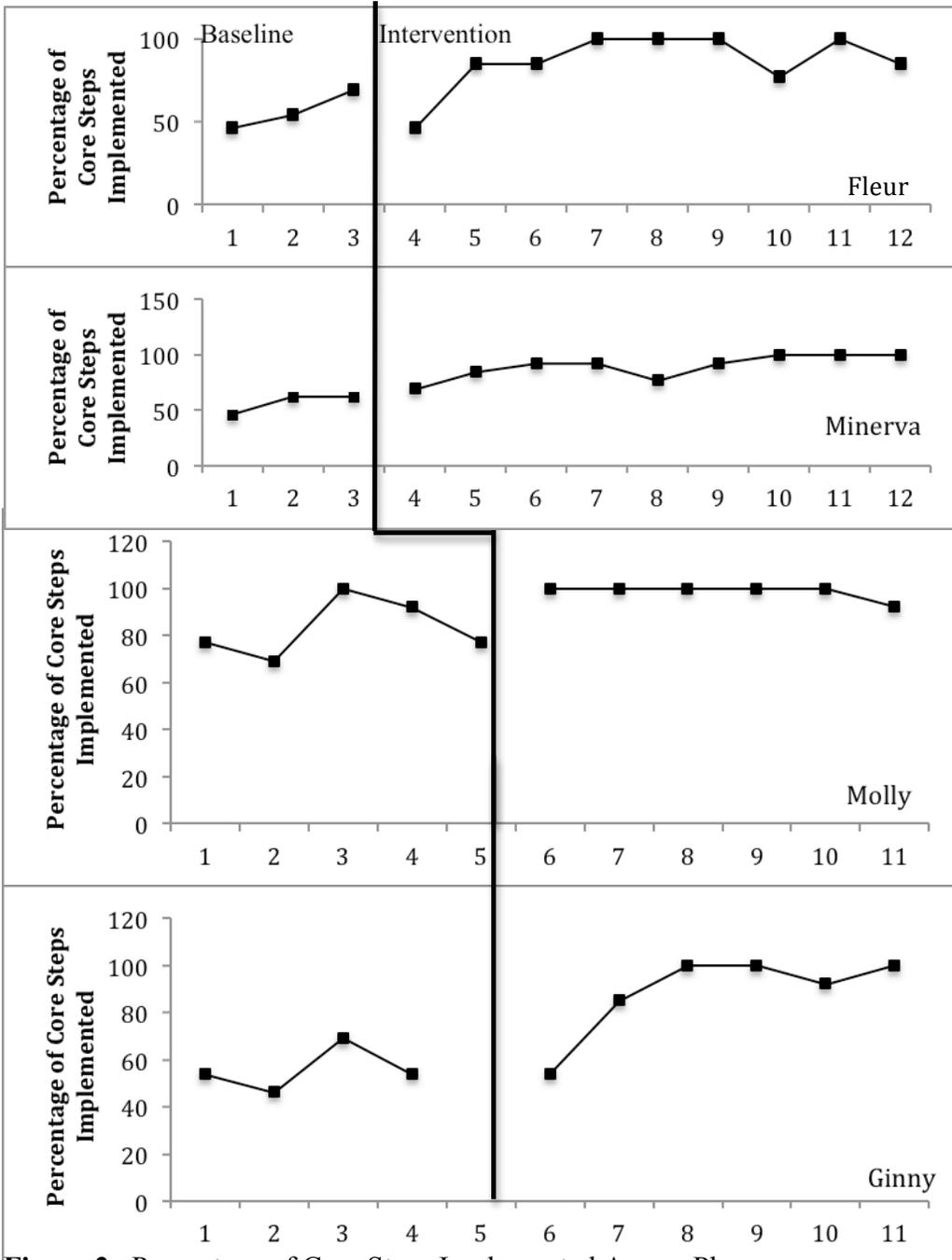


Figure 2. Percentage of Core Steps Implemented Across Phases

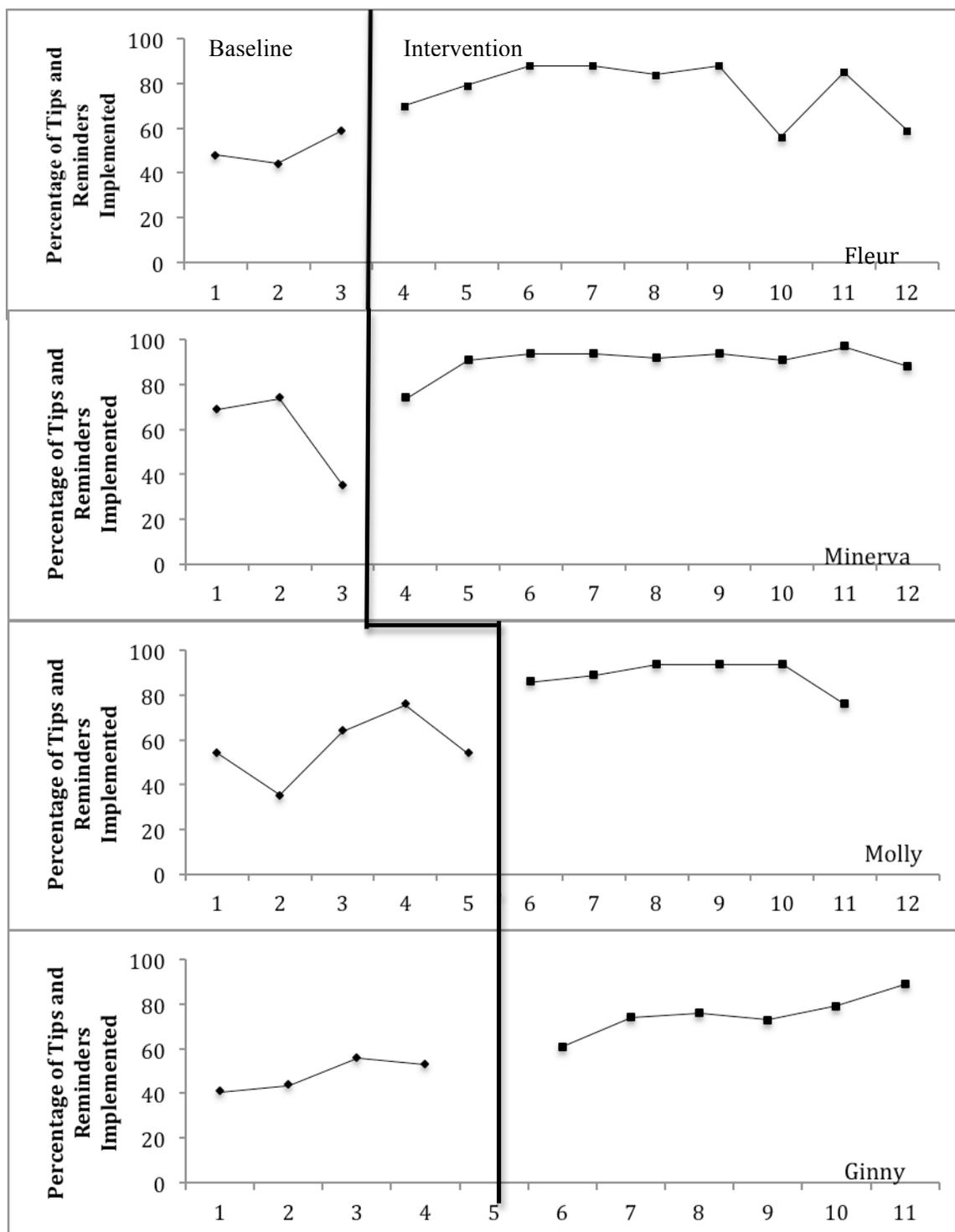


Figure 3. Percentage of Tips and Reminders Implemented Across Phases

Adherence to intervention steps. A cross-tabulation analysis was conducted per individual interventionist and across interventionists in order to determine if individual variability in interventionist's performance in the baseline and performance feedback phases predicted performance on steps. A cross-tabulation is a joint frequency distribution of cases that can be analyzed with the Goodman and Kruskal's Tau test in order to determine whether the variables are statistically independent or if they are associated (Field, 2006). The cross-tabulation acts as a frequency distribution, in that it predicts expected counts and compares them to the observed count. The differences between the expected and observed counts are then analyzed by the tau test to determine if one variable predicts another or not. A complete listing of percentage of implementation for Core Steps and Tips and Reminders for individual interventionists and across interventionists can be found in Tables 3 and 4. For a complete listing of steps, refer to Appendices B and C.

Analysis revealed that individual variability regarding performance on core steps at baseline was significant on the following Core Steps: *Phrase drill error correction* ($p < .05$), *Students read Passage aloud (3rd time)* ($p < .05$), and *Enthusiastic praise and feedback for students; distribution of stars based on Group Reading Goal and You/Me Game* ($p < .05$). In other words, individual variability between interventionists' performance predicted their adherence to the step. Individual variance did not significantly predict interventionist implementation of core steps during the feedback phase. Additionally, individual variability regarding performance significantly predicted step implementation on the following Tips and Reminders quality steps during the feedback phase: *Teacher specifically told students they met or did not meet the Group Reading Goal* ($p < .05$), *Teacher referred to the points that the*

*students earned as part of the You/Me Game ($p < .05$), Teacher referred to points that the students earned as part of the You/Me Game ($p < .05$), Teacher always stated the word that was read incorrectly before having the group read the phrase that contained the word ($p < .05$), Teacher asked students to practice “logical” phrases ($p < .05$), and With enthusiasm, teacher praised specific reading behaviors and praised group for specific reading behaviors or improvements at the end of the session ($p < .05$). Individual variance significantly predicted interventionist implementation of *step Before prompting student to begin the Retell Check*, teacher made sure students could not review the passage during the Retell Check ($p < .05$) of the Tips and Reminders quality steps at baseline only. It is important to note that there were six Tips and Reminders that could not be calculated because they are not always applicable (e.g., an interventionist would not circle a new data point unless the data point represented a new passage).*

Table 3 *Cross-tabulation of Core Step Implementation*

Interventionist	Phase	1	2	3	4	5	6	7	8	9	10	11	12	13
Fleur	Baseline	0	0	100	30	100	100	20	11.1	18.2	22.2	0	12.5	22.2
	Feedback	100	100	100	70	100	100	80	88.9	81.8	77.8	100	87.5	77.8
Minerva	Baseline	12.5	0	100	100	100	100	30	0	0	0	27.3	18.2	10
	Feedback	87.5	100	100	100	100	100	70	100	100	100	72.7	81.8	90
Molly	Baseline	25	25	100	45.5	100	40	45.5	40	100	100	40	45.5	37.5
	Feedback	75	75	100	54.5	100	60	54.5	60	100	100	60	54.5	62.5
Ginny	Baseline	20	25	44.4	33.3	100	33.3	28.6	33.3	33.3	100	16.7	25	16.7
	Feedback	80	75	55.6	66.7	100	66.7	71.4	66.7	66.7	100	83.3	75	83.3
Total	Baseline	13.8	12.5	34.1	33.3	100	30.2	31.6	22.9	26.3	28.2	22.9	26.3	21.2
	Feedback	86.2	87.5	65.9	66.7	100	69.8	68.4	77.1	73.7	71.8	77.1	73.7	78.8

Note. All numbers represent percentages.

Steps. 1. One-on-one assessment 2. Steps 1a-1d are completed with each remaining student in the group

3. Make decision about students meeting group reading goal 4. Review goals and you/me game 5.

Students read aloud 6. Passage retell 7. Model reading by teacher 8. Students read aloud 9. Phrase drill

error correction 10. Students read aloud 11. Praise and feedback/distribution of stars 12. Teacher records

group information on tracking form 13. Teacher reviews flow chart for possible missed steps and records this on tracking form.

Table 4 *Cross-tabulation of Tips and Reminders Implementation*

Interventionist	Phase	1	2	3	4	5	6
Fleur	Baseline	30	27.3	0	20	0	0
	Feedback	70	72.7	100	80	100	100
Minerva	Baseline	18.2	18.2	14.3	100	18.2	10
	Feedback	81.8	81.8	85.7	100	81.8	90
Molly	Baseline	40	25	14.3	100	28.6	40
	Feedback	60	75	85.7	100	71.4	60
Ginny	Baseline	40	25	33.3	100	25	28.6
	Feedback	60	75	66.7	100	75	71.4
Total	Baseline	31.7	23.7	16.7	32.6	17.6	21.2
	Feedback	68.3	76.3	83.3	67.4	82.4	78.8

Note. All numbers represent percentage of improvement

Steps. 1. Teacher had all materials available and organized before session 2. Teacher used scripted directions or abbreviated directions 3. Teacher specifically told group they met or did not meet their goal 4. Teacher always correctly indicated the number of words read in 1 minute 5. Teacher correctly scored and calculated each total WCPM and WIPM 6. While graphing teacher gave verbal feedback and praise.

Table 4

Interventionist	Phase	7	8	9	10
Fleur	Baseline	20	- ^a	0	11.1
	Feedback	80	-	100	88.9
Minerva	Baseline	30	-	0	20
	Feedback	70	-	100	80
Molly	Baseline	100	-	25	44.4
	Feedback	100	-	75	55.6
Ginny	Baseline	28.6	-	0	33.3
	Feedback	71.4	-	100	66.7
Total	Baseline	31.6	-	7.1	27
	Feedback	84.4	-	92.9	73

Note. All numbers represent percentage of improvement.

^a Dashes indicate lack of data due to steps being inapplicable at certain times of implementation.

Steps. 7. Teacher connected lines between WCPM and WIPM for the same passage 8. Teacher circled the data point and session number for new passages 9. Teacher completed each student's individual tracking form 10. Teacher randomly selected students to read.

Table 4 Continued

Interventionist	Phase	11	12	13	14	15
Fleur	Baseline	11.1	22.2	0	100	10
	Feedback	88.9	77.8	100	100	90
Minerva	Baseline	18.2	100	20	0	100
	Feedback	81.8	100	80	100	100
Molly	Baseline	45.5	37.5	50	16.7	33.3
	Feedback	54.5	62.5	50	83.3	66.7
Ginny	Baseline	40	33.3	50	33.3	25
	Feedback	60	66.7	50	66.7	75
Total	Baseline	29.3	28.9	22.2	20.6	23.1
	Feedback	70.7	71.1	77.8	79.4	76.9

Note. All numbers represent percentage of improvement.

Steps. 11. Teacher selected a new student to read once sentence was completed 12. Each student read 1-3 times 13. Teacher referenced points 14. Prompted closed books before retell 15. Teacher provided correct instructions for retell.

Table 4

Interventionist	Phase	16	17	18	19	20	21	22	23
Fleur	Baseline	100	- ^a	100	100	100	22.2	20	22.2
	Feedback	100	-	100	100	100	77.8	80	77.8
Minerva	Baseline	100	-	100	100	100	18.2	20	20
	Feedback	100	-	100	100	100	81.8	80	80
Molly	Baseline	40	-	40	100	100	33.3	14.3	20
	Feedback	60	-	60	100	100	66.7	85.7	80
Ginny	Baseline	100	-	100	100	100	16.7	0	0
	Feedback	100	-	100	100	100	83.3	100	0
Total	Baseline	31.8	-	31.8	100	100	22.9	15.6	20.8
	Feedback	68.2	-	68.2	100	100	77.1	84.4	79.2

Note. All numbers represent percentage of improvement.

^a Dashes indicate lack of data due to steps being inapplicable at certain times of implementation.

Steps. 16. Teacher called on each student to provide retell response 17. Teacher used broad follow up questions (if applicable) 18. Teacher read aloud at a pace just faster than group reading ability 19. Teacher read with good expression 20. Teacher read at an appropriate volume 21. Teacher paused 2-4 times 22. Teacher randomly selected students to read next word 23. Teacher referenced points

Table 4 Continued

Interventionist	Phase	24	25	26	27	28	29	30
Fleur	Baseline	18.2	16.7	28.6	18.2	0	18.2	- ^a
	Feedback	81.8	83.3	71.4	81.8	0	81.8	-
Minerva	Baseline	20	18.2	0	18.2	0	100	-
	Feedback	80	81.8	100	81.8	0	100	-
Molly	Baseline	28.6	40	0	100	0	45.5	-
	Feedback	83.3	60	100	100	0	54.5	-
Ginny	Baseline	0	33.3	50	100	100	44.4	-
	Feedback	100	66.7	50	100	0	55.6	-
Total	Baseline	20	27.8	15.8	30.2	100	28.2	-
	Feedback	80	72.2	84.2	69.8	0	71.8	-

Note. All numbers represent percentages of improvement.

^aDashes indicate lack of data due to steps being inapplicable at all times of implementation.

Steps. 24. Teacher always stated incorrect word before phrase reading 25. Teacher asked student to practice logical phrases 26. Teacher told students to “read” phrases, not “say” them 27. Teacher had students practice all missed words 28. Teacher pointed to words practiced 29. Teacher ensured students read phrases at the same time 30. Teacher told students to practice words read less fluently (if applicable)

Table 4 Continued

Interventionist	Phase	31	32	33	34	35	36	37	38
Fleur	Baseline	- ^a	0	0	0	0	- ^a	- ^a	0
	Feedback	-	100	100	100	100	-	-	100
Minerva	Baseline	-	0	18.2	0	18.2	-	-	18.2
	Feedback	-	100	81.8	100	81.8	-	-	81.8
Molly	Baseline	-	0	18.2	33.3	16.7	-	-	25
	Feedback	-	100	71.4	66.7	83.3	-	-	75
Ginny	Baseline	-	0	50	0	50	-	-	25
	Feedback	-	0	50	100	50	-	-	75
Total	Baseline	-	0	25	9.1	18.2	-	-	17.6
	Feedback	-	100	75	90.9	81.8	-	-	82.4

Note. All numbers represent percentages of improvement.

^aDashes indicate lack of data due to steps being inapplicable at all times of implementation.

Steps. 31. Teacher explained that words were correct but less fluent (if applicable) 32. Teacher referenced points 32. Teacher referenced points. 33. Teacher provided a minimum of three different praise statements regarding reading behavior 34. Teacher accurately told students why they earned each star and awarded no more than three 35. Teacher praised specific reading behaviors enthusiastically 36. If the group landed on or passed a shaded square on star chart they selected a ticket from the bonus bag 37. Teacher conveyed that improved reading skills, rather than the opportunity to earn stars is the primary reason to put forth effort (if applicable) 38. Teacher completed group tracking form with 100% accuracy.

Research Question Three

In order to determine differences in the amount of time it took for participants to reach high levels of treatment integrity compared to the study conducted by Begeny et al.

(2013), a comparison of means and medians were conducted. On average, participants in the

present study required 2.75 (*Mdn* = 3.5; *SD* = 2.4) sessions to reach high levels of treatment integrity on Core Steps versus 4.5 (*Mdn* = 4; *SD* = 2.2) sessions in the Begeny et al. (2013). However, comparisons could not be made between the current study and Begeny et al. (2013) because not all of the current interventionists met the stringent treatment integrity criteria posed in either study. In other words, half of the interventionists in the current study were able to implement the Tips and Reminders quality steps at 90% or higher in 3 of 5 consecutive sessions as demonstrated by participants in Begeny et al. (2013). Although they all at least met 80% TI on these steps, which has been deemed acceptable (DiGennaro, 2009), the current study implemented the same stringent criteria as Begeny et al. (2013) (i.e., 90% implementation or higher in 3 of 5 consecutive session) thus, reliable comparisons regarding the Tips and Reminders quality steps could not be made.

Survey Responses

Interventionists were given a five-question survey at the end of the study. Three of the four participants returned their surveys. In general, the three interventionists preferred verbal feedback to written feedback; used similar check systems to monitor what step they were on; and all found the individual readings, individual graphing, and group tracking the most difficult steps to implement. A breakdown of individual responses can be found in Table 5.

Table 5 *Interventionist Survey Responses*

Question	Interventionist Responses		
	Minerva	Molly	Ginny
Which type of feedback was most helpful (verbal, written, or both)? Why?	Both, but especially verbal. The day is so hectic that I often didn't take time to read the written feedback.	Written, I am just more visual.	Verbal most useful for me. I didn't read the written- sorry, no time.
Did you take additional steps to ensure that you were implementing all thirteen steps of the intervention each session (e.g., check off each step with a dry erase marker, practice at home, etc.)?	I reviewed it several times before I did the 1st session, and again each time before session for first 2 weeks. Once familiar, I just used checklist.	Just read over it a few times and had the sheet out each time I did small group.	Yes, I checked off each step with a marker.
Which steps of the intervention were easiest to implement?	-Reading aloud -Modeled Reading	Nothing was really difficult.	The whole group practicing.
Which steps of the intervention were the most challenging to implement? Why (e.g., easy to forget, not clearly defined, time restraints, etc.)?	-Earning of stars was confusing at first -1-on-1 is time consuming -Too many graphs -Graphs were confusing	Each day having to do individual readings to chart. Sometimes they had to read a few passages each and time was a problem.	Hardest for me was all the charting (group, individual, etc.). Lots to keep up with.
Is there anything else you would like to say about the feedback or intervention that hasn't been addressed?	Team members from NCSU were positive and helpful.	No Response.	I enjoyed working with NC State. Than you for you patience!

Discussion

The purpose of the current study was to measure, monitor, and ultimately increase the treatment integrity of a small-group reading fluency intervention using combined written and verbal performance feedback procedures. The results regarding interventionist treatment integrity indicated that although all teachers improved significantly on their implementation of HELPS-SG Core Steps and Tips and Reminders, there were still marked drops in treatment integrity after reaching 100% implementation. The following discussion will focus on the contextual factors affecting the study, implications of the results, limitations, and future research directions.

Contextual Factors

As an applied research project in a real world setting, the results of this study are considered to demonstrate external validity. For example, this study was conducted within an already stressed framework, which is likely a common depiction of many schools in a large district. The district and the schools in which the research took place were rolling out new programs, assessments, and various other changes and although it is not uncommon for school districts to roll out new initiatives, these changes likely influenced results in a variety of ways. For instance, the training workshop was originally intended for a professional development day and to last approximately three hours. However, the school, along with the rest of the district, experienced great technical difficulties with a new report card system and had to hold an emergency meeting regarding grade distribution on the day of the interventionists' training. As a result, the workshop was reduced to about an hour. The

researcher then met with interventionists per grade level to complete training prior to the start of the study.

Additionally, the frequent and not uncommon occurrence of school events such as field trips, assemblies, half days, and progress monitoring, often disrupted the planned schedule. In order to conduct research within the pre-existing school schedule, readjusting the research schedule was necessary, resulting in a reduction of the total number of sessions that were originally intended.

A final note on school context concerns teacher acceptance of the intervention and participation in the study. In addition to aforementioned externally mandated changes, the principal recruited teachers for the current study. Although two teachers volunteered freely, two others were asked by the principal to participate. This combination of free choice and possible coercion raised some trepidation regarding teacher acceptance, as they were expected to implement these new mandates with integrity as well as participate in a highly demanding research project. Teacher concerns related to acceptance were first noted after the training workshop. Given that the current study took place within the school's regular schedule and all interventionists demonstrated moderate to significant increases in their TI, it can be assumed that the implemented performance feedback procedures were effective and this study demonstrated high external validity (DiGennaro et al., 2005; Erchul & Martens, 2010).

Discussion of Individual Cases

Fleur. A cross-tabulation of the steps revealed Fleur's improvement was on Core Steps such as the individual assessments (e.g., graphing and correctly choosing the appropriate passage(s) to read) as well as not skipping certain steps, such as the second group practice reading or the group choral reading. Regarding her predicted implementation of steps as analyzed by the Goodman-Kruskal Tau, Fleur's lack of adherence at baseline to step 11 (i.e., praise and distribution of stars at end of session) contributing the most variability to predicted adherence as compared to the other interventionists. Other notable variance in her adherence to steps included her difficulty in consistently implementing Tips and Reminders quality steps after feedback, as indicated by the Goodman-Kruskal Tau analysis. Fleur did not consistently specify to the group if they had met their reading goals, state the word that was being read incorrectly before practicing with phrase drill exercises, or choose logical phrases to practice. Other areas of inconsistency, as highlighted per the Goodman-Kruskal Tau, were seen among all interventionists on the Tips and Reminders quality steps, namely, steps related to the motivational procedures. For instance, Fleur inconsistently reminded students of the points from the "You/Me" game at various points where it is listed in the abbreviated and scripted directions. Fleur did demonstrate gains regarding her implementation of Tips and Reminders step 24, stating the incorrect word before practice during the phrase drill step. Her implementation increased from 66% at baseline to 100% after feedback. Of note, Fleur demonstrated substantial variability regarding star distribution and providing enthusiastic praise at the end of the session. Her implementation of Tips and Reminders step 35 increased from 0% at baseline only to 33% at feedback. In Fleur's case,

stars typically exceeded the maximum amount of three and were inconsistently awarded. As an augmentation to the reward system, Fleur attempted to link the amount of stars earned during the group related to a monetary reward system that she used in her classroom. Again, given the earlier concern regarding acceptability, it was not unexpected that she failed to meet the high criteria set for the Tips and Reminders. However, despite these challenges, it is important to note that she demonstrated moderate improvement on her implementation of the Tips and Reminders.

Fleur demonstrated two drops in TI on Core Steps after she had reached 100% implementation and never met the high criteria set for the Tips and Reminders. Important to note is the marked drop in TI after her first feedback session (i.e., Fleur implemented HELPS-SG only 46% correctly after day one of feedback). However, during this session, she made multiple attempts at correcting mistakes. Unfortunately, she struggled to make corrections smoothly; in fact, the more she realized she was implementing steps out of order or incorrectly, the more flustered she became and unable to correct herself. Although her implementation of steps remained relatively low, it is important to note that even after one feedback session, Fleur became more aware of her implementation of the intervention.

Toward the end of the study, Fleur experienced two drops in TI after reaching 100% implementation of the Core Steps. There are two hypotheses advanced as to these drops in TI: reactance and acceptability (Briesch, Chafouleas, Rak Neugebauer, & Riley-Tillman, 2013; Noell, 2008; Wilkinson, 2006). Reactance is an expected limitation of direct observations methods (e.g., Cochrane & Laux, 2008; Lane et al., 2004; Noell, 2008). The researcher purposefully had a small team of assistants who were assigned to the same

teachers daily in order to reduce reactance and build rapport. Toward the end of the study, however, Fleur's assigned observer was not available and had to be replaced with a new observer. This observer noted that Fleur's hands were shaking, and Fleur herself noted in the audio recording that she "always graphed before starting the group session," and "never had to check through the steps before," both of which reflect incorrect implementation of steps. Fleur had previously implemented the referenced steps correctly, thus reactance is another factor that might explain her TI drops and her defensive attitude toward a new observer.

Acceptability is another possible explanation related to Fleur's observed drops in TI. Emerging research in implementation science suggests that popular theories from the 1980s indicating that teacher acceptance is a one-dimensional construct and the sole link to treatment integrity fails to consider the multifaceted nature of acceptance (Briesch et al., 2013; Dart, Cook, Collins, Gresham, & Chenier, 2012). In contrast, developing literature proposes that teacher acceptance or acceptability of an intervention has come to encompass three levels of influence, which typically lead to treatment usage (immediate and sustained) and treatment integrity (Briesch et al., 2013; Mautone, DuPaul, Jitendra, Tresco, Vile Junod, & Volpe, 2009). In other words, teacher acceptability requires more than just the teacher accepting an intervention as valid.

The first level or component of acceptability is individualistic. If an individual does not perceive the intervention to be fair or appropriate, he or she likely will not use it. The second level is consideration of the feasibility of the proposed intervention or treatment; even a highly effective intervention is less likely to be implemented if it creates a significant disruption to the environment where something easily adaptable will likely be implemented.

The final level includes contextual and environmental considerations. More specifically, administrative and building-level supports also influence intervention acceptability and usage. Thus, even when teachers feel competent using the intervention and approve of its use, physical and administrative support are essential to achieving higher levels of acceptability (Briesch et al., 2013; Mautone et al., 2009).

Fleur was one of the teachers asked by the school principal to participate in the study and she related to the researcher that she only agreed so that she could add the experience toward her professional development goals and growth. Throughout the study, Fleur commented that she looked forward to the end of the study and never wanted to participate in any other research because it was so time consuming. Additionally, she often asked to reschedule her assigned times and cancelled twice at the start of the study due to field trips. Thus, acceptability is another possible explanation related to her drops in TI. Given the aforementioned circumstances, it is important to note that feedback effects still remained moderately effective for Fleur.

Minerva. A cross-tabulation of steps revealed that Minerva significantly improved on almost all Core Steps. Of particular importance were improvements in individual assessments (e.g., graphing and completing individual tracking charts), the second and third group practice readings, and the choral phrase practice reading. As one of two teachers who met the high criteria at one point for the Tips and Reminders, she showed great improvement in her overall accuracy in graphing and tracking of progress and the motivational aspects of the intervention (e.g., correctly giving stars, correctly tracking points for the You/Me game,

and specific praise). Additionally, she was one of two teachers to reach the stringent Tips and Reminders criteria of 3 of 5 consecutive sessions at 90% implementation or higher.

As demonstrated by the Goodman-Kruskal Tau analysis, at baseline Minerva never adhered to the phrase drill step and never conducted the third student reading of the passage, which contributed the most variance related to adherence when compared to other interventionists. For instance, throughout the baseline phase she did not adhere to Tips and Reminders step 14: prompting students to close their books before answering questions as part of the retell step. The Goodman-Kruskal Tau also revealed some variability in Minerva's performance on motivational procedures. For example, she inconsistently reminded students of the points they had earned at various intervals when implementing the "You/Me" game and unreliably awarded the correct amount of stars, even after feedback. However, variance in her performance at feedback on Tips and Reminders quality steps was not as great as compared to other interventionists.

Minerva did not experience any drops in TI once she reached 100% implementation on the Core Steps. She also volunteered freely and was the teacher most interested in the intervention. She was very active during her feedback sessions, asking many questions before and after, including asking for modeling of missed steps once the sessions were over. Similarly, Minerva consistently reviewed the Flow Chart and Scripted Directions before starting each session. She also related to the researcher and her undergraduate observer that she felt motivated by the observed increases in her students' fluency during sessions. It was therefore not surprising that the PND statistic revealed strong effects for her feedback procedures (cf. Briesch et al., 2013; Mautone et al., 2009).

Molly. A cross-tabulation allowed the researcher to understand improvements on Molly's specific steps pre-and post-feedback. On Core Steps, Molly improved upon her overall implementation of individual assessments (e.g., accurately choosing the correct passage(s) to read). As the second of two teachers who met the high criteria at one point for the Tips and Reminders, Molly demonstrated marked improvement in her overall accuracy in tracking both individual and group information and random assignment of students to read sections of the passage being practiced.

As demonstrated by the Goodman-Kruskal Tau analysis, at baseline Molly was inconsistent when awarding stars and providing praise as related to Core Step 11. Additionally at baseline, Molly inconsistently adhered to Tips and Reminders step 3; specifically tell students whether or not they have met their group reading goal before beginning the session. However, after feedback she adhered to this step 100% of the time. Similarly, she inconsistently adhered to Tips and Reminders step 14; prompting students to close their book before answering questions related to the retell step. After feedback her variability in adhering to this step remained. The Goodman-Kruskal Tau also revealed some variability in Molly's performance on motivational procedures. For instance, she inconsistently reminded students of the points they had earned at various times during the "You/Me" game and inconsistently awarded the correct amount of stars, even after feedback. At times, Molly would dismiss the group before awarding them stars. Finally, Molly consistently did not offer specific praise or feedback regarding reading behaviors at either baseline or feedback phases. In fact, her performance on this Tips and Reminders quality step decreased after feedback, and added the most variance regarding adherence to the step.

Molly experienced one drop in TI after she had reached 100% implementation and is the only other teacher to have met the stringent Tips and Reminders criteria (although she did experience a drop in the quality steps after reaching criteria). It is hypothesized that this drop was a result of a change of schedule and reactance (i.e., Cochrane & Laux, 2008; Lane et al., 2004; Noell, 2008). Molly's last session (i.e., the session that fell below 100% implementation) had to be rescheduled to a day that the intervention normally did not take place. Furthermore, her usual observer (the researcher) could not attend the session and had to be replaced by a research assistant with whom Molly had not yet worked. An audio recording of the session revealed that Molly felt "a little flustered" and apologized for rushing that day. The research assistant reported that Molly appeared uncomfortable during the session and was rushing. Although Molly's PND was not deemed valid due to extreme data points, visual analysis indicated that feedback did have some positive impact on her implementation of Core Steps, as she again reached 100% implementation after her first feedback session. Moreover, there were moderate effects of feedback regarding the Tips and Reminders, which can be seen in her ability to reach the stringent a priori criteria (cf. Briesch et al., 2013; Mautone et al., 2009).

Ginny. Regarding implementation of specific steps, Ginny demonstrated improvement in her implementation of choosing the correct passage to read during individual assessments and correctly tracking group information at the end of each session. Most impressively, her PND statistic of 100% on her implementation of the Tips and Reminders quality steps indicated that the performance feedback procedures were highly effective. In fact, she came close to reaching 90% implementation during her last session. It could be

hypothesized that, had the research project lasted longer, she too would have reached the high a priori criteria.

Although she demonstrated marked improvements, a closer look at the frequency counts of the cross-tabulation and the Goodman-Kruskal Tau analysis of the Core Steps Tips and Reminders quality steps revealed similar inconsistencies to all interventionists in recognizing whether or not the group met their reading goal, providing specific praise, reminding students of the points earned in the “You/Me” game, and incorrect awarding of stars to students.

Regarding her implementation of Core Steps, the Goodman-Kruskal Tau analysis revealed inconsistent performance on the phrase drill procedure, having students read the passage a third time, and praise, feedback and star distribution. Of note, Ginny inconsistently adhered to the phrase drill procedures at baseline; however, implemented the step 100% of the time after feedback. Similarly, she adhered to the third round of passage reading only 50% of the time at baseline and improved to 100% implementation after feedback. Conversely, Ginny remained inconsistent in her adherence to Core Step 11, praise, feedback, and star distribution at both baseline and feedback. In fact, her performance on this step decreased after feedback, as did Fleur’s.

Regarding her performance on the Tips and Reminders quality steps per the Goodman-Kruskal Tau analysis, Ginny was the only interventionist to improve her implementation to 100% after feedback on step 3, specifically telling students whether or not they met their group reading goal. Similarly, she improved to 100% implementation of prompting students to close their books before retell began after feedback. Ginny

demonstrated the most difficulty adhering to Tips and Reminders step 23;referring to points earned during the You/Me game. In fact, even after feedback, Ginny never adhered to this step, which is stated in the abbreviated and scripted directions. She was the only interventionist that did not improve her performance on this step, contributing to the overall variance in the Goodman-Kruskal Tau analysis. Finally, Ginny's performance on Tips and Reminders step 35, specific praise regarding reading behaviors, was unreliable. Similarly to other interventionists, she inconsistently praised students at both baseline and feedback. Again, she remained the only interventionist that demonstrated a decrease in her implementation of this step at feedback.

Although the principal recruited Ginny, acceptance was not a concern particularly once the feedback procedures began. She experienced one drop in TI after she reached 100% Implementation of the Core Steps and never met the high criteria set for the implementation of the quality steps, Tips and Reminders. It is hypothesized that this drop was also a result of observer reactance (cf. Cochrane & Laux, 2008; Lane et al., 2004; Noell, 2008). The researcher was assigned to Ginny for the duration of the study, but Ginny had to reschedule her last day to a time the researcher was unavailable. The researcher had built strong rapport with Ginny, and it should be noted that Ginny was frequently anxious about making errors, often commenting "I can't wait until you can just tell me how bad I'm doing everything, I just wish you could help me now" during the baseline phase. A research assistant worked with Ginny during her final session, and the audio recording of the session revealed that Ginny noted when she made a mistake and became upset with herself because "I had been getting it right!" The assistant also noted that Ginny appeared rushed and unprepared for the

session at the start. That being noted, Ginny still demonstrated marked improvements in both her implementation of the Core Steps and her implementation of the Tips and Reminders quality steps.

Survey Results

All interventionists were asked to complete a brief survey at the conclusion of the study. As previously noted from the three interventionists who did return surveys, most comments reflected observational data surrounding the idea of doing more with less. The interventionist who did not return the survey was the same one who raised the greatest concerns regarding teacher acceptance. Written responses to the survey questions were all brief, at times only consisting of one-word answers. Thus, observational data coupled with survey answers may reveal more regarding their perceptions surrounding the current study and the intervention.

All interventionists who responded commented on the feedback procedures throughout the study. Most maintained that they did not always have time for the verbal consultations and often felt rushed to get through them. As part of a larger and ongoing project surrounding teacher stress, Stauffer and Mason (2013) found that 91% of their teacher participants ($n = 64$) indicated comparable recurring stressors including workload and time to accomplish tasks and teaching. Additionally, participants cited concerns regarding the amount of time spent on paperwork and not feeling informed or prepared for changes in policy and procedures (Stauffer & Mason, 2013). It might be hypothesized that although most teachers indicated that the feedback they received was helpful, they did not feel that

feedback was a priority as it was an addition to their curriculum and instructional time. However, there is also research that suggests feedback may be perceived as disciplinary or negative (DiGennaro et al., 2005). Thus, it might also be hypothesized that for some participants, the feedback and observations were perceived as punitive once all TI requirements had been met (DiGennaro et al., 2005).

Participants in the current study referenced both the amount of time needed to implement the intervention and the amount of time spent completing all tracking forms. The fastest overall time of implementation (including the individual assessment component) was 15 minutes. This time, however, does not include the additional time it took to complete all of the necessary tracking forms or the additional feedback session. It may be hypothesized that the entire implementation process was longer than what was recorded.

Overall implementation time for all interventionists ranged from 15-21 minutes post-feedback. This range is a marked improvement from the baseline phase, where complete implementation including individual assessments ranged from 24-40 minutes (HELPS-SG Flow Chart). Although there was no formal manual available for HELPS-SG during the time this study was conducted, the Flow Chart does offer approximations for time regarding both the individual assessment and the group assessment. It is estimated that one individual assessment takes between 1 and 2 minutes and the group assessment takes between 7 and 12 minutes. However, these times are not inclusive of the data tracking; rather, these times reflect only the time it takes to implement the reading program. A survey of 1,000 teachers (Primary Sources, 2012) revealed that teachers want more time to help their students in need, yet typically allocate only about 20 minutes of their instructional time during the day to do

so. Similarly, teachers reported working an average of approximately 10 hours and 40 minutes daily on instruction, including tutoring, parent outreach, and specific interventions. This amount is over the required amount of time that teachers reportedly are required to be present at school (Primary Sources, 2012). Finally, the survey indicated that teachers do want more professional development time in order to address student needs but do not often receive the support or have the time (Primary Sources, 2012; Stauffer & Mason, 2013). The idea of doing more with less, particularly regarding time, was a recurring theme in participants' responses in the current study. Regular education teachers are not necessarily afforded the amount of time it likely takes to implement HELPS-SG or a similar type of intervention without some kind of additional support. Given these contextual insights, interventionists' remarks regarding time represent a valid concern.

Despite issues regarding acceptance and time, it is also clear the performance feedback procedures in the current study were effective. However, there were limitations to the study that can help guide future research in the area. These limitations and future directions are discussed in the following section.

Limitations and Future Directions

The most important limitations of this study were participant attrition and lack of time. First, a participant attrition rate of 50% made it difficult reach definitive conclusions. Although it is typical of single-subject designs to have only four participants (e.g., Dart et al., 2012; DiGennaro et al., 2007; Kaufman et al., 2013), this study demonstrated the time-intensive nature of direct methods of observation and TI research. Second, time was a

significant factor that influenced results. Interventionists were expected to implement 15-20 sessions as opposed to the 11 or 12 that actually occurred. Adjustments made to the schedule resulted in a shortened baseline phase for the first group of teachers (i.e., 3 sessions), the inclusion of the proposed maintenance phase, and an inability to conduct statistical analyses. In other words, a Wilcoxon signed-rank test could not be conducted as originally proposed due to lack of data; therefore, the researcher could not truly assess changes across phases.

Another significant limitation of this study was participant selection. Half of the teachers were chosen by the school principal in contrast to the original proposal to use current intervention teachers who would volunteer freely. This result is consistent with extant literature regarding teacher acceptance and teacher choice (Anderson & Daly, 2013). Specifically, Anderson and Daly (2013) found that teachers who do not feel supported by their administration tend to perceive an intervention as less acceptable. Likewise, when teachers are given a choice to “test-drive” or choose interventions, they are more likely to implement them with high integrity and use the intervention consistently in their classrooms (Anderson & Daly, 2013). Relative to the present study, the originally proposed participant population of intervention teachers had implemented the one-on-one version of the HELPS program, and it was hypothesized that their prior knowledge would compromise the design of the current study. Consequently, the researcher used general education classroom teachers. Originally eight teachers were selected; however, only six participated in the workshop and two dropped out after the study began. This significant drop in participation created concern regarding teacher acceptance for the researcher, as the literature supports an association between treatment integrity and treatment acceptability (Mautone et al., 2009). Three

quarters of the remaining interventionists were interested in the intervention, which likely increased their acceptance of the overall research project. After rapport was built, all but one teacher demonstrated a reasonable level of acceptance. It should be noted though, that despite challenges related to teacher acceptance, the performance feedback measures improved all interventionists' implementation. This is a promising discovery.

The interventionists' training was another limitation. As previously mentioned, this training was intended to last approximately three hours; however, due to an emergency meeting, the workshop lasted only an hour. Although the researcher met with interventionists by grade level to complete training, they were not afforded the opportunity to practice in groups led by the researcher. These shortened instructional sessions may account for the interventionists' inability to meet the high criteria set for the quality steps, Tips and Reminders, as the researcher had less time at the training to review these steps. Additionally, the intervention that was implemented did not yet have a manual, which made self-training more difficult. Unlike the traditional one-on-one HELPS program (Begeny, 2009), which includes a manual and video vignettes of each step, the only materials available to the interventionists were the Flow Chart, Scripted and Abbreviated Directions, and the Tips and Reminders.

Similarly, the researcher encountered many scheduling conflicts, as one would expect given that the project took place in a school. At the start of the study, the schedule fluctuated greatly per individual interventionist. At times, the interventionists completed both the individual and group sessions at once, although most times the individual assessments occurred in the morning and the group sessions were conducted in the afternoon. The

intervention had to be implemented three times a week, ideally every other day (e.g., Monday, Wednesday, Friday); however, due to Fridays having a different schedule than the rest of the week, the sessions were held on Mondays, Wednesdays, and Thursdays. Even when provided with this change, this schedule had to be adjusted for each interventionist almost every week. Additionally, two interventionists missed one session each due to curriculum-based measurements and mid-year assessments.

Another limitation to this study is its inability to examine differences between verbal feedback versus written feedback, in that only the combined effects of both were studied. In a similar vein, only one type of written feedback was provided. Current literature (e.g., Zoder-Martell, Dufrene, Sterling, Tingstrom, Blaze, Duncan, & Harpole, 2013) suggests that graphical or visual data regarding TI is another useful form of written feedback. The current study only provided the Tips and Reminders checklist to interventionists as their written feedback.

Finally, it is possible that the quality of the feedback and/or the quality of the rapport between observer and interventionist moderated interventionists' improvement. These possible moderation effects were not accounted for in the current study, nor has it been previously accounted for in the extant TI literature.

Given the aforementioned limitations, possible future research directions might aim to examine the individual effects of verbal feedback versus written feedback. Recently, Kaufman et al. (2013) attempted to discover the differences between verbal feedback only and written feedback only. They found that although both types were effective in increasing TI, verbal feedback yielded increases in TI more quickly than written feedback. However,

they did not employ the use of graphical or visual TI data and teachers were not afforded the opportunity to ask questions about the written feedback they were given. Kaufman et al. indicated that these limitations warrant further research. Perhaps in an email form, teachers could have the opportunity to ask questions in a written format as opposed to waiting for a consultative meeting (i.e., verbal feedback phase) to ask questions.

It is also important to know to what effect different forms of written feedback have on TI. For instance, if graphical or visual information of interventionists' TI is more effective than a checklist, then classroom consultants should adjust their written feedback to interventionists. Additionally, if the combined effect of emailing checklists, graphs, and verbal feedback is most effective, then classroom consultants and other faculty or staff should adjust their feedback methods to increase TI. This approach might also be less demanding on interventionists' time yet still help improve overall TI.

Additionally, more research should be done to examine moderating effects of feedback quality. The current study used highly trained observers with specialized knowledge about the HELPS-SG program. In a school setting, not all interventionists or consultants have the opportunity to become experts in all interventions that are implemented in a classroom. Therefore, it is important to better understand to what extent the quality of feedback (e.g., highly specific and positive versus general or broad statements) affects interventionists' implementation levels.

Finally, single-subject designs provide the ability to observe and understand the settings in order to provide a richer interpretation of contextual information. Most interventionists were concerned about the limited amount of time they had and being

expected to do more with less. It may be hypothesized that if teachers had more help in the classroom in the form of trained teachers aides or volunteers, they would have been more open to the research process and, in the case of this study, even continued the HELPS-SG intervention once the researcher left. Most TI literature focuses on the process of implementation and less on who delivers the intervention (e.g., Bradshaw et al., 2009; DiGennaro et al., 2007; Noell & Gansle, 2014; Solomon, Klein, & Politylo, 2012). Perhaps TI research should begin focusing on the implementation integrity levels of volunteers in order to determine if this population can implement interventions with acceptable integrity (Begeny et al., 2013; Mattingly, Prislin, McKenzie, Rodriguez, & Kayzar, 2002). One emerging organization, the Metropolitan Organization of Volunteers Empowering Students (MOVES), is under development and based on the idea of training student volunteers in a number of interventions and connecting them with schools in need.

Overall Conclusions and Implications

Currently, many schools employ a multi-tiered framework (e.g., response to intervention; RtI) that uses a problem-solving approach in order to address varying levels of student need based on universal screening and targets those students with gaps between expected and actual performance (Reschly & Bergstrom, 2009). As part of the RtI movement, small group interventions are increasingly used, particularly within the second tier, and are seen as an integral solution in maximizing already limited school resources (Wasik, 2008). However, with the increased use of small group EBIs increased attention should be focused on the TI of the interventionists. This is particularly important for classroom consultants to monitor, as RtI is an alternative framework for making eligibility

decisions for special education as compared to a traditional test and place model (Kratochwill et al., 2007). Finally, a growing body of literature suggests that performance feedback is the best method to use when increasing TI. However, issues regarding teacher acceptance and observer reactance remain challenges that must be overcome by both teachers and classroom consultants (Briesch et al., 2013; Cochrane & Laux, 2008; Lane et al. 2004; Noell & Gansle, 2014).

The current study aimed to monitor, measure, and increase the treatment integrity of a small-group reading fluency intervention using performance feedback measures. The applied setting allowed for a better understanding of how interventions are implemented in schools and what kinds of roadblocks may exist. Despite issues revolving around acceptance and the relatively complicated and time consuming nature of the intervention used, the combined effect of verbal and written performance feedback both rapidly and effectively increased all interventionists' levels of implementation.

References

- Barnett, D., Hawkins, R., & Lentz, F.E. (2011). Intervention adherence for research and practice: Necessity or triage outcome? *Journal of Educational and Psychological Consultation, 21*, 175-190.
- Begeny, J.C., Upright, J.J., Easton, J.E., Ehrenbock, C.A., & Tunstall, K.R. (2013). Validity estimates and functionality of materials and procedures used to monitor the implementation integrity of a reading intervention. *Journal of Applied School Psychology, 29*, 284-304.
- Begeny, J. C. (2009). *Helping Early Literacy with Practice Strategies (HELPS): A one-on-one program designed to improve students' reading fluency*. Raleigh, NC: The Helps Education Fund. Retrieved from <http://www.helpsprogram.org>.
- Berkeley, S., Bender, W.N., Peaster, L.G., & Saunders, L. (2009). Implementation of response to intervention: A snapshot of progress. *Journal of Learning Disabilities, 42*, 85-95.
- Bradshaw, C.P., Debnam, K., Koth, C.W., & Leaf, P. (2009). Preliminary validation of the implementation phases inventory for assessing fidelity of school wide positive behavior supports. *Journal of Positive Behavior Interventions, 11*, 145-160.
- Briesch, A.M., Chafouleas, S.M., Rak Neugebauer, S., & Riley-Tillman, T.C. (2013). Assessing influences on intervention implementation: Revision of the usage rating profile-intervention. *Journal of School Psychology, 51*, 81-96.

- Cochrane, W.S., & Laux, J.M. (2008). A survey investigating school psychologists' measurement of treatment integrity in school-based interventions and their beliefs about its importance. *Psychology in the Schools, 45*, 499-507.
- Dart, E.H., Cook, C.R., Collins, T.A., Gresham, F.M., & Chenier, J.S. (2012). Test driving interventions to increase treatment integrity and student outcomes. *School Psychology Review, 41*, 467-481.
- DiGennaro, F.D., Martens, B.K., & McIntyre, L.L. (2005). Increasing treatment integrity through negative reinforcement: Effects on teacher and student behavior. *School Psychology Review, 34*, 220-231.
- DiGennaro, F.D., Martens, B.K., & Kleinmann, A.E. (2007). A comparison of performance feedback procedures on teachers' treatment implementation integrity and students' inappropriate behavior in special education classrooms. *Journal of Applied Behavior Analysis, 40*, 447-461.
- Erchul, W. P., & Martens, B. K. (2010). *School consultation: Conceptual and empirical bases of practice*. (3rd ed.). New York, NY: Springer.
- Foorman, B.R., & Torgesen, J. (2001). Critical elements of classroom and small-group instruction promote reading success in all children. *Learning Disabilities Research & Practice, 16*, 203-212.
- Fuchs, D., & Fuchs, L.S. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly, 41*, 93-99.

- Fuchs, D., & Fuchs, L.S. (2009). Responsiveness to intervention: Multilevel assessment and instruction as early intervention and disability identification. *The Reading Teacher*, 63, 250-252.
- Good, R. H., Kaminski, R. A., Cummings, K., Dufour-Martel, C., Petersen, K., Powell-Smith, K., Stollar, S., & Wallin, J. (2011). *Dynamic indicators of basic early literacy next assessment manual*. Eugene, OR: Dynamic Measurement Group, Inc.
- Retrievable from <http://www.dibels.org>.
- Gresham, F.M., Macmillan, D.L., Beebe-Frankenberger, M.E., & Bocian, K.M. (2000). Treatment integrity in learning disabilities intervention research: Do we really know how treatments are implemented? *Learning Disabilities Research & Practice*, 15, 198-205.
- Hughes, C.A., & Dexter, D.D. (2011). Response to intervention: A research-based summary. *Theory into Practice*, 50, 4-11.
- Individuals with Disabilities Education Improvement Act of 2004 (IDEIA), Pub. L. No. 108-446 Stat. 2647 (2004).
- Jimerson, S.R., Burns, M.K., & VanDerHeyden, A.M. (2007). Response to intervention at school: The science and practice of assessment and intervention. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to intervention: The science and practice of assessment and intervention* (pp. 3-10). New York, NY: Springer.
- Kaminski, R., & Good, R.H. (2004). *Dynamic Indicators of Basic Early Literacy Skills*. Multi-Health Systems, Inc.

- Kaufman, D., Coddling, R.S., Markus, K.A., Shick Tryon, G., & Nagler Kyse, E. (2013). Effects of verbal and written performance feedback on treatment adherence: Practical application of two delivery formats. *Journal of Educational and Psychological Consultation, 23*, 264-299.
- Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings*. (2nd ed.). New York, NY: Oxford University Press.
- Kratochwill, T.R., Clements, M.A., & Kalymon, K.M. (2007). Response to intervention: Conceptual and methodological issues in implementation. In S.R Jimerson,, M.K. Burns, & A.M. VanDerHeyden, (Eds.), *Handbook of response to intervention: The science and practice of assessment and intervention* (pp. 25-52). New York, NY: Springer.
- Lane, K.L., Bocian, K.M., MacMillan, D.L., & Gresham, F.M. (2004). Treatment integrity: An essential – but often forgotten – component of school-based interventions. *Preventing School Failure, 48*, 36-43.
- Lenz, A.S. (2012). Calculating effect size in single-case research: A comparison of nonoverlap methods. *Measurement and Evaluation in Counseling and Development, 46*, 64-73.
- Mautone, J.A., DuPaul, G.J., Jitendra, A.K., Tresco, K. E., Vile Junod, R., & Volpe, R.J. (2009). The relationship between treatment integrity and acceptability of reading interventions for children with attention-deficit/hyperactivity disorder. *Psychology in the Schools, 46*, 919-931.

- Mitchell, B.S., Stormont, M., & Gage, N.A. (2011). Tier two interventions implemented within the context of a tiered prevention framework. *Behavioral Disorders, 36* (4), 241-261.
- National Center on Response to Intervention. (2010, 04). *Essential components of rti - a closer look at response to intervention*. Retrieved from www.rti4success.org/pdf/rtiessentialcomponents_042710.pdf.
- No Child Left Behind, 20 U.S.C. § 16301 *et seq.* (2001).
- Noell, G.H. (2008). Research examining the relationships among consultation process, treatment integrity, and outcomes. In W.P. Erchul & S.M. Sheridan (Eds.), *Handbook of research in school consultation* (pp. 323-341). New York: Lawrence Erlbaum Associates/Taylor & Francis Group.
- Noell, G.H., & Gansle, K.A. (2014). Research examining the relationships between consultation procedures, treatment integrity, and outcomes. In W.P. Erchul & S.M. Sheridan (Eds.), *Handbook of research in school consultation* (2nd ed.; pp. 386-408). New York, NY: Routledge.
- Perepletchikova, F., Treat, T.A., & Kazdin, A.E. (2007). Treatment integrity in psychotherapy research: Analysis of the studies and examination of the associated factors. *Journal of Consulting and Clinical Psychology, 75*, 829-841.
- Primary sources: 2012 america's teachers on the teaching profession. (2012, January 1). Retrieved from http://mediaroom.scholastic.com/files/ps_fullreport.pdf.
- Reschly, D. J., & Bergstrom, M. K. (2009). Response to intervention. In T. Gutkin & C.

- Reynolds (Eds.), *The handbook of school psychology* (4th ed.) (pp. 434-460).
Hoboken, NJ: John Wiley & Sons.
- Sanetti, L.M., & Kratochwill, T.R. (2009). Toward developing a science of treatment integrity: Introduction to the special series. *School Psychology Review*, *38*, 445- 459.
- Schulte, A.C., Easton, J. & Parker, J. (2009). Advances in treatment integrity research: Multidisciplinary perspectives on the conceptualization, measurement, and enhancement of treatment integrity. *School Psychology Review*, *38*, 460-475.
- Scruggs, T.E., & Mastropieri, M.A. (1998). Summarizing single-subject research: Issues and applications. *Behavior Modification*, *22*, 221-242.
- Solomon, B.G., Klein, S.A., & Politylo, B.C. (2012). The effect of performance feedback on teachers' treatment integrity: A meta-analysis of the single-case literature. *School Psychology Review*, *41*, 160-175.
- Stage, S. A, Jackson, H.G., Moscovitz, K., Jensen Erickson, M., Ogier Thurman, S., Jessee, W., & Olson, E.M. (2006). Using multimethod-multisource functional behavioral assessment for students with behavioral disabilities. *School Psychology Review*, *35*, 451-471.
- Stauffer, S.D., Mason, E.C.M. (2013). Addressing elementary school teachers' professional stressors: Practical suggestions for schools and administrators. *Educational Administration Quarterly*, *49*, 809-837.
- Torgesen, J., Wagner, R., & Rashotte, C. (2012). *Test of Word Reading Efficiency* (2nd ed.). Austin, TX: Pro-Ed.

- Wasik, B. (2008). When fewer is more: Small groups in early childhood classrooms. *Early Childhood Education Journal*, 35, 515-521.
- Wiederholt, J.L., & Bryant, B.R. (2012). *Gray Oral Reading Tests* (5th ed.). Austin, TX: Pro-Ed.
- Wilkinson, L.A. (2006). Monitoring treatment integrity: An alternative to the 'consult and hope' strategy in school-based behavioural consultation. *School Psychology International*, 27, 426-438.
- Zoder-Martell, K., Dufrene, B., Sterling, H., Tingstrom, D., Blaze, J., Duncan, N., & Harpole, L. (2013). Effects of verbal and graphed feedback on treatment integrity. *Journal of Applied School Psychology*, 29, 328-349.

APPENDICES

Appendix A:
Helping Early Literacy with Practice Skills - Small Group reading intervention (HELPS-SG)
Instructional Materials

HELPS Program for Small Groups: Implementation Protocol

Steps below that are denoted with an asterisk should be implemented in combination with the “HELPS Program for Small Groups: Scripted Directions.”

One-on-One Assessment Procedure:

1. One student at a time is asked to read passages aloud to the teacher (*1.5 to 2.5 minutes per student*).
 - a. *The student reads the passage practiced in the most previous session (Passage A).
 - b. *The student reads the next passage in the sequenced curriculum (Passage B).

Note: Step 1b is skipped if (a) the student read Passage B in the most previous one-on-one assessment session; and (b) the group used Passage A (not Passage B) during the most previous session, which occurs when the group fails to meet the Group Reading Goal in the previous session.
 - c. Teacher records the student’s WCPM and WIPM (for each passage) on that student’s Graph and Individual Progress Tracking Form.
 - d. Teacher provides specific feedback and enthusiastic praise for reading improvements.
2. Teacher implements Steps 1a – 1d with each remaining student from the group.

Group-based Instructional Strategies:

3. Teacher determines whether at least half of the students met the Individual Reading Goal (IRG). The IRG is determined by each student’s grade level, as shown in the Table below.

Individual Reading Goals (IRGs) According to Each Student’s Grade Level

	WCPM with Passage A	WIPM with Passage A
First Grade	80 or more	3 or less
Second Grade	100 or more	3 or less
Third Grade	120 or more	3 or less
Fourth Grade	135 or more	3 or less

The IRGs above are suggested Goals for students at the given grade level. As described in the Teacher’s Manual, it is sometimes appropriate for the teacher to modify a student’s IRG if necessary.

Group Reading Goal (GRG): The Criterion and Instructional Decision Rules

The GRG criterion is met when at least half of the group’s students meet his/her IRG. (*0.25 minutes*)

- a. If the GRG criterion is met, teacher praises the students and tells them they will (a) earn 1 star on the Group Star Chart, and (b) practice a new story during that session (Passage B).
 - b. If the GRG criterion is not met, teacher verbally acknowledges students’ improvements but tells them they will again practice the story read during the previous session (Passage A).
4. Teacher reads introductory statements and expectations to the group (*0.5 minutes*)
 - a. *Teacher says, “As a group, you’re going to be doing some reading with me today. As you read, I want you to do your best reading. This means I want you to read as quickly as you can without making mistakes, and try to read with good expression (like I do when I read to you). I also want you to remember what happens in the story and try to remember the difficult words that we practice.”
 - b. *Teacher reminds students about the You/Me Game. If the students follow along well enough to know the place in the story when called on by the teacher, the group earns a point. If a student is called on and he/she does not know the place in the story, the teacher earns a point. If the students beat the teacher by the pre-determined amount of points and read with full effort, they earn 1 to 2 additional stars on the Group Star Chart. (See Teacher’s Manual for recommendations about selecting a pre-determined amount of points).
 5. Group Reading 1: Students read the selected passage (Passage A or B) aloud (*1.5 to 2 minutes*)

- a. *Teacher provides directions for the group read-aloud.
 - b. One randomly selected student starts reading aloud by reading the first 1 to 3 sentences of the passage.
 - c. All other students are instructed to follow along silently and wait for their turn to read aloud.
 - d. After the first student finishes reading, the teacher randomly selects another student in the group to read the next 1 to 3 sentences.
 - e. This procedure continues until the students read for approximately two minutes or until the passage is complete.
 - f. If a student is called on, he/she must immediately start reading where the previous student left off. If the student does this, the team earns a point in the You/Me Game. If the student does not, the teacher earns a point in the You/Me Game.
6. *Teacher asks students to retell what they remember about the story, asking students to retell what happened in appropriate sequential order *(0.5 to 1 minute)*
 - a. One student is randomly called on to say one thing that he/she remembers from the story.
 - b. This procedure is repeated until all students have been called on once.
 7. *Teacher models fluent oral reading of passage while each student follows along with his/her finger *(1 to 2 minutes)*
 - a. Periodically throughout the reading, the teacher calls on a student to say the next word in the passage.
 - b. If a student is called on, he/she must immediately say the next word in the passage. If the student does this, the team earns a point in the You/Me Game. If the student does not, the teacher earns a point in the You/Me Game.
 8. Group Reading 2: Students read the session's passage aloud a second time *(1.5 minutes)*
 - a. Teacher implements Steps 5a – 5f.
 9. *Teacher implements phrase-drill error correction on all incorrectly read words from Group Reading 2 *(0.25 to 0.75 minutes)*
 - a. All students chorally respond to phrases the teacher asks them to practice.
 - b. After all phrases are practiced, the teacher assigns 0 – 5 points for the You/Me Game based on collective student effort.
 - c. If time permits (i.e., very few words were read incorrectly in Group Reading 2), complete the phrase-drill procedure with words read incorrectly in Group Reading 1 and/or practice otherwise difficult words or phrases from the passage.
 10. Group Reading 3: Students read the session's passage aloud a third time *(1.5 minutes)*
 - a. Teacher implements Steps 5a – 5f.
 11. Teacher provides enthusiastic praise and feedback on the students' reading and praises effort as applicable *(0.5 to 1 minute)*
 - a. If students met the GRG criterion, teacher adds 1 star to the Group Star Chart and states the reason the students earned that star.
 - b. Based on the outcome of the You/Me Game, teacher tells students they earned an additional 1 or 2 stars on the Group Star Chart.
 - c. Students select from bonus bag and/or prize box, as applicable.
 12. Teacher records group information on the Group Progress Tracking Form.
 13. Teacher reviews implementation steps from flow chart and records any steps missed on the Group Progress Tracking Form.

Total implementation time: 7.5 to 10.5 minutes (excluding one-on-one assessment procedure)

Appendix B:
Helping with Early Literacy Practice Skills-Small Group reading intervention (HELPS-SG)
Implementation Materials

HELPS Program for Small Groups: Scripted Directions

(For one-on-one assessments) Directions to administer before student reads passage:

1. Place the teacher copy of the reading passage in front of you but shielded so the student cannot see what you record. Present the student copy of the reading passage to the student.
2. Say to the student, **“Here is a story that I would like you to read. When I say ‘Begin’, start reading aloud at the top of the page and read across the page. Try to read each word. If you come to a word you don’t know, I’ll tell it to you. Do you have any questions? Be sure to read as quickly as you can without making errors, and try to read with good expression.”**
3. Say, **“Begin!”** and start the stopwatch when the student says the first word.
4. If the student hesitates on a word for more than 3 seconds, say the word.
5. At the end of one minute, place a closed bracket after the last word.
6. *If the student reads so fast that no expression is given, remind the student that when he/she reads the next story, you want him/her to read at a comfortable rate (i.e., with good expression).*
7. Remove both copies of the reading passage and record student’s score on the Individual Progress Tracking Form.

Introductory statements and expectations (includes Verbal Cuing Procedure):

As a group, you’re going to be doing some reading with me today. As you read, I want you to do your best reading. This means I want you to read as quickly as you can without making mistakes, and try to read with good expression (like I do when I read to you). I also want you to remember what happens in the story and try to remember the difficult words that we practice.

Describing the You/Me Game:

As we read, we will play the You/Me Game. To earn points for your team and win the game, you need to work hard and follow directions throughout today’s lesson. For example, when I am reading aloud or another student is reading aloud, you need to read along silently to yourself. If I call on you and you know where we are in the story, your group will earn one point. If you do not know where we are in the story, I will earn one point. At the end of the lesson, if your team earns XX [insert # based on recommendations for the Teacher’s Manual] more points than I do, your group will earn two stars on your Group Star Chart.

Directions to administer before students read passage aloud (Group Reading Procedure):

1. Say to the group, **“Now we are going to practice reading today’s story out loud. (Insert student’s name) will start reading when I say *begin*, and then I will call on another student to read. Each of you will take turns reading a few sentences at a**

time. When you are not reading aloud, you need to read along silently while your classmate reads. If I call on you and you know where we are in the story, your group will earn one point in the You/Me Game. If you do not know where we are, I will earn one point.”

2. Say “**Begin**” and have the designated student start reading aloud. After this student reads 1 to 3 sentences, randomly select another student in the group to read the next 1 to 3 sentences.
3. Continue this procedure of randomly selecting students to read 1 to 3 sentences until the group reads for approximately two minutes or until the students read the entire passage.
4. When calling on a new student to read, if that student immediately starts reading where the previous student left off, record one point for “You” (the students) in the You/Me Game. If the student cannot start where the previous student left off, record a point for “Me” (the teacher).
5. As students read, record students’ reading errors on the teacher copy of the passage. Record errors with slashes, underlines, or circles, as instructed for Reading 1, 2, or 3.
6. At the end of the activity, briefly praise the group for their effort (as applicable).

Directions for administering Retell procedure:

1. Say to the group, “**Now I would like each of you to tell me one important part that you remember from the story. Try to tell me what happened in the correct order.**”
 - For expository text, you should say, “**Now I would like each of you to tell me one important thing that you learned from the reading. Starting with the first important thing you learned, tell me about what you learned in the correct order.**”
2. Randomly select one student to begin the Retell procedure. When that student finishes a brief (approximately 10 seconds) retell, randomly select another student to retell a portion of the passage.
3. Continue this procedure until all students have been called on once. If necessary, give prompts to help students remember key parts of the passage and to correctly sequence the events.
4. At the end of the Retell procedure, briefly praise the group for their effort (as applicable).

Directions for teacher to read passage aloud (Modeling Procedure):

1. Say to the group, “**Now I am going to read today’s story to you. Please follow along with your finger, reading the words to yourself as I read them. Sometimes I will stop reading to make sure you are following along. If I call your name, you need to tell me the next word in the story. If you read the correct word, this will show me you are reading along with me and your group will earn 1 point in the You/Me Game. If you do not show me you are reading along with me, I will earn one point.**”

2. Read the passage at a comfortable reading rate and with good expression for approximately 1.5 minutes or until you read the entire passage. Make sure the students are following along with their fingers and prompt students to do this, if necessary.
3. While reading the passage, stop 5-7 times in order to randomly call on a student to read the word that immediately follows the word you stopped at. Record a point for the group (a “You” point in the You/Me Game) when a student reads the correct word. Record a point for the teacher (a “Me” point) when a student cannot immediately read the next word.
7. At the end of the activity, briefly praise the group for their effort (as applicable).

Directions for administering Phrase-drill Error-correction Procedure:

1. Say to the students, “**Now we are going to practice some of the difficult words from the story.**” Have the students sit closely enough so they can all see the one passage you will point to.
2. On a student’s copy of the passage, point to the first error word from Group Reading 2, say the word, and then say, “**All together, read this after I do,** [read the 2-5 word phrase containing the error word and then have the students chorally read the phrase]. **Again** [students should chorally read the phrase a second time], **Again** [students should chorally read the phrase a second time].” This procedure should allow the students to chorally read the phrase three times.
 - a. To better ensure all students read the phrase at the exact same time, use a prompt (e.g., snap your finger, tap the back-end of a marker on the table) to signal for students to begin reading the phrase aloud.
3. Make sure students read the phrases rather than simply memorize them and repeat them. (Teachers want students to read, rather than recite). Also be sure that all students read aloud together.
4. Repeat the above procedure for all *unique* error words from Group Reading 2. If time permits (i.e., very few words were read incorrectly in Group Reading 2), complete the Phrase-drill procedure with words read incorrectly in Group Reading 1.
 - a. If students make 1 or fewer errors total on Readings 1 and 2, practice 1-3 words or phrases the students read less fluently. Use the procedures above.
5. After all phrases are practiced, assign 0-5 points for the You/Me Game, based on collective student effort. Praise students’ effort accordingly.

HELPS Program for Small Groups Implementation Flow Chart

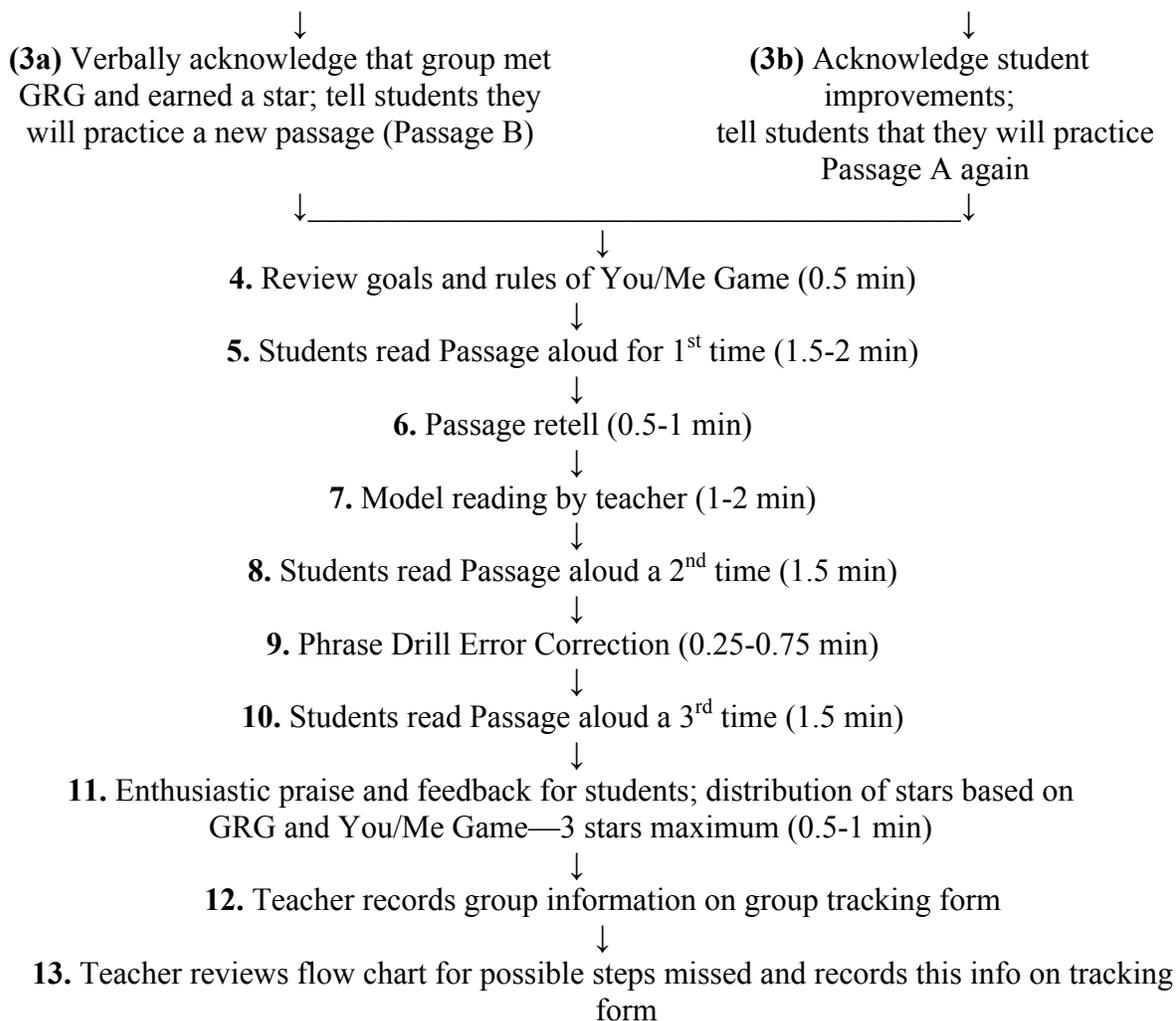
1. One-on-one assessment with each student in group, facilitated by any trained educator

- a. ORF assessment with Passage A
- b. ORF assessment with Passage B (as needed)
- c. Record data on Graph and Tracking Form
- d. Offer praise and feedback

2. Steps 1a – 1d are completed with each remaining student from the group.

3. Make decision about students meeting the Group Reading Goal (GRG)...

At least ½ of the students meet _____ OR _____ Less than half of the students meet the IRG (see table below)



Individual Reading Goals (IRGs) According to Each Student's Grade Level

	WCPM with Passage A	WIPM with Passage A
First Grade	80 or more	3 or less
Second Grade	100 or more	3 or less
Third Grade	120 or more	3 or less
Fourth Grade	135 or more	3 or less

Appendix C:
Materials to Monitor Treatment Integrity using the Helping Early Literacy with Practice
Skills -Small Group reading intervention (HELPS-SG)

Steps and Guidelines for Observing Teachers during their Implementation of the HELPS-Small Group Program

- 1. On the Observation Summary Form (OSF), record (a) your name, (b) the teacher's name, (c) the date, (d) the observation (OBS) session number (e.g., if the teacher has been observed 4 times previously, write "5" because this is the 5th OBS session), (e) the name of student/group receiving HELPS-SG, and (f) the group HELPS-SG session number.
- 2. Wait patiently until the teacher is ready to begin implementing HELPS-SG with the student and observe the teacher's organization and preparation for the session.
- 3. When the teacher begins Step 1 of HELPS-SG implementation, start your stopwatch to begin monitoring the total time it takes the teacher to complete the session.
- 4. Throughout the HELPS-SG implementation session, record all steps the teacher completes correctly and incorrectly on the Observation Checklist for Implementing the HELPS-SG Program. Steps should be recorded for implementation of both Core Procedures and Tips and Reminders.
- 5. At the end of the entire HELPS-SG session (after the students return to class and the teacher completes the Progress Tracking form), stop your stopwatch and record the following information on the OSF: (a) whether the group met their goal, (b) the duration of the HELPS-SG session in minutes and seconds, (c) the number or % of steps the teacher completed accurately from the Core Procedures checklist, (d) the Step #s (e.g., 3, 5a) not implemented from the Core Procedures (if applicable), (e) the number or % of steps the teacher completed accurately from the Tips/Reminders checklist, (f) your evaluation of the teacher's enthusiasm during the session, and (g) your evaluation of the teacher's organization during the session.
- 6. Reset and start your stopwatch again. If there were steps from the Core Procedures checklist the teacher did not implement, review those missed steps with the teacher and discuss any questions the teacher has about those missed steps. When appropriate, provide a rationale for why a step should be performed in a particular way. (**Applicable?**)
- 7. Next, identify **at least** 2-3 steps that the teacher carried out correctly. The praise that you provide should be **genuine, enthusiastic, and specific** (e.g., "Nice job accurately describing why the student earned the star on his chart by telling him he earned it for meeting his goal on the first reading"). When applicable, you should provide specific feedback about "targeted improvements" the teacher made since an earlier OBS session (related to step 14 below).
- 8. Share the data you recorded in parts *b, c, d, e, f,* and *g* (from step 7 above) with the teacher.
- 9. If there were steps from the Tips and Reminders checklist the teacher did not implement, review those missed steps with the teacher and, if needed, discuss any questions the teacher

- has about those missed steps. **(Applicable?)**
- 10. If you did not rate the teacher's enthusiasm and/or organization as "outstanding," provide a rationale of your evaluation to the teacher and discuss the situation as needed. **(Applicable?)**
 - 11. If applicable, provide additional feedback (i.e., feedback not related to the Implementation Protocol or Tips and Reminders Checklist) to the teacher that will likely help him/her implement HELPS-SG better in the future. For example, you may offer advice about how to organize HELPS-SG materials in the most effective and time efficient way. **(Applicable?)**
 - 12. Ask the teacher if he/she has any questions about HELPS-SG implementation procedures (or the program, in general) and answer/discuss those questions as needed.
 - 13. Record the topics discussed in steps 10 and/or 11 on your OSF.
 - 14. At the end of the OBS session, identify 1-3 things (as deemed appropriate/applicable) the teacher should improve upon during subsequent HELPS-SG sessions with students (these are considered "targeted steps" for the teacher to improve). Make sure the teacher has a final opportunity to ask questions about what to improve and how to do so. You should demonstrate the 1-3 step(s) as needed. Note: the 1-3 targeted steps for improvement should have already been discussed in Steps 11, 12, 13, and/or 14 above. **(Applicable?)**
 - 15. Thank the teacher for his/her time and effort and conclude the OBS session. **Overall, the teacher should finish each OBS session feeling positive and better prepared to implement the HELPS-SG Program, rather than feeling judged or deemed inadequate.**
 - 16. Stop your stopwatch and record on your OSF: (a) the duration of the post-session OBS meeting, (b) whether all teacher questions/concerns were addressed, (c) whether all missed steps and tips/reminders were reviewed, (d) all Core Procedures not implemented, and (e) any additional, meaningful notes about the meeting. Finally, based on the 1-3 targeted steps to improve (described in step 14 above), specify these targeted steps on your OSF or write "None" if no Core Procedures or Tips/Reminders were missed.
 - 17. Make a copy of the HELPS-SG Observation Checklist and provide it to the teacher immediately after your feedback discussion.
 - 18. Review steps 1-17 above and: place a "✓" in the box for completed steps; place a circle "O" around the box for non-applicable (NA) steps; place a "X" in the box for skipped steps; and use arrows "←→" to specify if and how a step was implemented out of order. This self-feedback should help you to reduce or eliminate Xs or arrows in your next observation session with a teacher.

Percentage of steps completed = total steps completed / total steps applicable: ___/___x100 = ___%

Examples of Major (versus Minor) Procedural Errors for each Core HELPS-SG Instructional Procedure

Core Instructional Procedure	Examples of Major Implementation Errors
Verbal Cueing procedure (i.e., the introductory statement and expectations)	<ul style="list-style-type: none"> Teacher does not provide any portion of the introductory statement to student and/or group before the student reads first passage of session (individual and/or group session)
Individual Assessment	<ul style="list-style-type: none"> Teacher does not obtain a cold read after the first retention of Passage A
Repeated Reading (Timed Reading) procedure	<ul style="list-style-type: none"> Teacher does not provide any form of directions or provides inaccurate directions before starting procedure Teacher has student and/or group read the wrong passage
Retell procedure	<ul style="list-style-type: none"> Teacher does not provide any form of directions or provides inaccurate directions before starting procedure
Modeling procedure	<ul style="list-style-type: none"> Teacher does not provide any form of directions or provides inaccurate directions before starting procedure Teacher never pauses to have a student in the group read next word in the passage
Phrase-drill Error Correction procedure	<ul style="list-style-type: none"> Teacher does not provide any form of directions or provides inaccurate directions before starting procedure Teacher does not have group practice WIPM (as recorded by teacher in most previous Timed Reading)—applicable only when student has WIPM
Goal Setting procedure	<ul style="list-style-type: none"> Based on teacher's recording of individual student data, teacher incorrectly determines whether student met the Individual Reading Goal by assessing the WCPM, WIPM, and Retell Check criteria Teacher incorrectly determines whether the group met the reading goal based on all students' Individual Reading Goal Teacher states that group met Reading Goal but does not start group on next passage in Step 5
Performance Feedback/Graphing procedure	<ul style="list-style-type: none"> While graphing the individual student's performance, teacher does not show student the graph at all before dismissing student

Motivational (Reward) Procedure	<ul style="list-style-type: none">• Teacher provided no praise during entire session• Teacher never specifies points that are earned by students in the You/Me Game• Teacher told group that they earned a different number of stars than they actually earned• Teacher does not acknowledge that student earned a prize from prize box—applicable only when student receives a star in the last square of a shaded
---------------------------------	--

Observation Summary Form (OSF) for HELPS-Small Group Program

Teacher's Name: _____

Observer's Name: _____

Date: _____ Observation (OBS) session #: _____

Second Observer (if applicable): _____

Group receiving HELPS: _____ Group's session #: ____ [Group met goal: Yes No] [Session Duration: _____]

(Rate Enthusiasm and Organization/preparation 1 – 5; 1=poor, 3=average, 5=outstanding)

Teacher implemented _____ % of primary protocol, and _____ % of tips/reminders. [Enthusiasm with student: ____] [Organization: ____]

Core Procedures not implemented (List step numbers and write notes if needed): _____

Observer notes during OBS (e.g., notes of tips/reminders not implemented):

Questions or concerns raised by teacher (or additional notes or feedback provided by observer): _____

Targeted Steps:

Duration of OBS session: _____ [All teacher questions/concerns were addressed: Yes No] [All missed steps & tips were reviewed: Yes No]

Teacher's Name: _____

Observer's Name: _____

Date: _____ Observation (OBS) session #: _____

Second Observer (if applicable): _____

Group receiving HELPS: _____ Group's session #: ____ [Group met goal: Yes No] [Session Duration: _____]

(Rate Enthusiasm and Organization/preparation 1 – 5; 1=poor, 3=average, 5=outstanding)

Teacher implemented _____ % of primary protocol, and _____ % of tips/reminders. [Enthusiasm with student: ____] [Organization: ____]

Core Procedures not implemented (List step numbers and write notes if needed): _____

Observer notes during OBS (e.g., notes of tips/reminders not implemented):

Questions or concerns raised by teacher (or additional notes or feedback provided by observer): _____

____ **Targeted Steps:**

Duration of OBS session: _____ [All teacher questions/concerns were addressed: Yes No] [All missed steps & tips were reviewed: Yes No]

Observation Checklist for Implementing the HELPS-Small Group Program

Teacher observed: _____ Observer: _____ Group: _____
Date: _____

Implementation of Individual Assessment

(Place “✓” in the box for completed steps; Place “X” in the box for skipped steps or those implemented with major errors; use arrows “←→” to specify if and how a step was implemented out of order).

<u>List of Steps</u>				
<u>Student</u>	<u>1a</u>	<u>1b</u>	<u>1c</u>	<u>1d</u>
1.				
2.				
3.				
4.				
5.				

1. Steps 1a-1d above were completed correctly with at least one student
2. Steps 1a-1d above were completed correctly with ALL students in the group
3. Teacher made correct decision about students meeting the Group Reading Goal

Implementation of Core Group Procedures

List of Core Group Steps: 4 ; 5 ; 6 ; 7 ; 8 ; 9 ; 10 ; 11 ; 12 ; 13

Percentage of steps completed = total steps completed / total steps possible ___/13 x 100: ___%

Implementation of Tips and Reminders (steps missed below are considered minor errors)
(Place “✓” in the box for completed steps, Place circle “O” around the box for non-applicable [NA] steps)

General Implementation Procedure

- Teacher had the following materials available and organized before starting the session: stop watch, examiner passage, student passage, dry-erase marker, pencil, student graph, Progress Tracking Form (for the group and each individual), Star Chart, Bonus Bag, Implementation Flow Chart, and Scripted Directions. Also, the prize box was reasonably accessible.

- Teacher used Scripted Directions or Abbreviated Directions as advised at top of Abbreviated Directions.
- Teacher specifically told students they met or did not meet the Group Reading Goal before starting Step 5.

Timed Reading, Feedback, and Graphing Procedures during Individual Assessments

- After each student's oral reading, teacher always correctly indicated on the examiner passage (with a bracket) the number of words read in one minute.
- Teacher correctly scored and calculated each student's total WCPM and WIPM per passage read.
- While graphing, teacher gave verbal feedback and praise regarding the student's WCPM and WIPM scores.
- Teacher connected lines between WCPM (and WIPM) scores *only* for scores of the same passage.
- Teacher circled the data point and session number when the group began a new passage. **(Applicable?)**
- Teacher completed each student's Individual Progress Tracking Form with 100% accuracy.

Group-based Repeated Readings

- Teacher randomly selected students in the group to read (i.e., no consistent order was used to select students to read).
- Teacher selected a new student to read only when the prior student completed a sentence.
- Each student read 1-3 sentences when it was his or her turn to read.
- Teacher referred to points that the students earned as part of the You/Me Game.

Retell Check Procedure

- Before prompting student to begin the Retell Check, teacher made sure students could not review the passage during the Retell Check.
- Teacher provided the correct directions for Retell Check depending on whether the students read a narrative or expository passage.
- Teacher called on each student in the group to provide a response.
- Teacher used broad follow-up questions to solicit a student's retell *only* if the student was unable to retell anything for approximately 10 seconds. **(Applicable?)**

Modeling Procedure

- Teacher read aloud at a pace just a little faster than the group's reading ability.
- Teacher read with good expression.
- Teacher read at a volume the group could clearly hear.
- Teacher paused 5-7 times to have a student read the next word in the passage.
- Teacher randomly selected which student would read the next word in the passage (i.e., no consistent order was used to select students to read).
- Teacher referred to points that the students earned as part of the You/Me Game.

Phrase-Drill Error Correction Procedure

- Teacher always stated the word that was read incorrectly before having the group read the phrase that contained the word.
- Teacher asked students to practice “logical” phrases.
- Teacher told students to “READ” the phrases, and did not ask them to “SAY” or “REPEAT” phrases.
- Teacher had students practice all incorrectly read words from group reading #2 (Step 8), up to 5 words or until time permitted)
- Teacher pointed (or had the students point) to each word practiced.
- Teacher ensured that all students read the phrases at the same time and (if needed) used a prompt to (e.g., snap of finger, tap of marker on the table) to signal for students to begin reading the phrase aloud.
- If students made 1 or fewer errors, the teacher told the student(s) to practice 1-3 words or phrases that were read less fluently or were reading incorrectly in a previous student reading. **(Applicable?)**
- If students practiced words that were read *correctly* but less fluently (see above step), teacher explained to the student(s) that they read the words correctly, but will practice them because they are difficult. **(Applicable?)**
- Teacher referred to points that the students earned as part of the You/Me Game.

Motivational (Reward) Procedure and Data Recording

- Throughout the session, teacher provided a minimum of three different praise statements regarding group’s reading behavior.
- When awarding stars on the Group Star Chart, teacher accurately told students why they earned each star (Group Reading Goal and/or You/Me Game, up to 3 stars total)
- With enthusiasm, teacher praised *specific* reading behaviors (e.g., nice job reading *accurately* and with *good expression*; I like how you *corrected words you missed previously*) and praised group for specific reading behaviors or improvements at the end of the session.
- If the group landed on OR passed a shaded square on Group Star Chart, they were allowed to select a ticket from the bonus bag and teacher correctly recorded the bonus stars written on the ticket. **(Applicable?)**
- Teacher conveyed that improved reading skills, rather than the opportunity to earn stars/prizes, is the primary reason the group should put forth effort during each HELPS session. **(Applicable?)**
- Teacher completed the Group Progress Tracking form with 100% accuracy.

Total steps applicable = 38 total check boxes – number of boxes circled as NA ____ =

Total steps completed = number of boxes with a check mark = _____

Percentage of items completed = total items completed / total items applicable ___ / ___ x
100

Percentage of items completed: _____ %

Inter-scorer Reliability Agreement (ISRA) of each Student's Timed Readings

Discrepancies / Total words read:

Reading 1: ___ / ___ ISRA%: _____

Reading 2: ___ / ___ ISRA%: _____

Reading 3: ___ / ___ ISRA%: _____

Reading 4: ___ / ___ ISRA%: _____

Time of Session (TOS): _____

Appendix D: Teacher Consent Form

Teacher Consent Form

Title of Study: Evaluating Treatment Integrity and Elementary Student Outcomes of the HELPS-Small Group Reading Fluency Intervention Program – A Two-Part Study

Principal Investigators:

Faculty Sponsor:

What are some general things you should know about research studies?

You are being asked to take part in a research study. Your participation in this study is voluntary. You have the right to be a part of this study, to choose not to participate, or to stop participating at any time without penalty. The purpose of research studies is to gain a better understanding of a certain topic or issue. You are not guaranteed any personal benefits from being in a study. Research studies also may pose risks to those that participate. In this consent form, you will find specific details about the research in which you are being asked to participate. If you do not understand something in this form, it is your right to ask the researcher for clarification or more information. A copy of this consent form will be provided to you. If at any time you have questions about your participation, do not hesitate to contact the researchers named above.

What is the purpose of this study?

The overall purpose of this two-part study is to measure teachers' implementation of the HELPS-SG reading fluency intervention as well as to measure student outcomes after participating in the HELPS-SG reading intervention.

The purpose of part one of the study is to measure teachers' implementation of HELPS-SG, and provide them with individual performance feedback regarding their implementation.

The purpose of part two of the study is to measure students' reading performance after receiving HELPS-SG reading fluency intervention. The following question will be addressed in this study: Will students who receive HELPS-SG significantly improve on reading performance measures more so than expected?

What will happen if you take part in the study?

If you agree to participate in this study, you will be asked to attend a HELPS-SG training workshop. After, you will be asked to attend one-on-one meetings to receive feedback and review your performance implementing the HELPS-SG reading intervention. You will be asked to participate in these meetings until a predetermined cut-off criteria of 100% accuracy on core implementation steps and 90% accuracy on supplemental steps.

The primary investigator for part one of the study will make copies of the observation checklists at the end of each session and provide you with a copy.

During part two of the study, you will implement HELPS-SG 3 times a week with predetermined groups of 3-4 of your students who are struggling readers. While implementing HELPS-SG, a research assistant from North Carolina State University will observe 20-30% of the HELPS-SG sessions to ensure you are maintaining proper implementation protocol. They will not participate in your proceedings, and will act as a “fly-on-the-wall” during these sessions. At two time points (prior to starting implementation of HELPS-SG, and after implementing 40 sessions of HELPS-SG), research assistants will individually assess your participating students’ reading performance with commonly used reading assessments. Once data collection is completed, the primary investigator of part two of the study will collect information about how many HELPS-SG sessions each student participated in.

Risks

Potential risks for the first part of the study include the knowledge of observations and potential confidentiality of teachers. This will be addressed in the following ways. First, observers will work within teachers’ normal reading group schedules to minimize attention to those teachers being observed. Additionally, all participating teachers will be assigned a number for all forms, to ensure that all personal identification is confidential. Teachers will be de-identified at the conclusion of each session by the principal investigators, at their research laboratory at North Carolina State University. This information, along with any and all copies of written feedback will be kept according to APA guidelines, in a locked office on North Carolina State University. Finally, any and all feedback regarding teacher implementation will not be shared with school administration or be used in any way as a means for school administration to evaluate job performance.

Benefits

There is a strong need to support implementation of interventions in schools and to know how much a student is receiving planned intervention steps in order to make data-based decisions. Furthermore, it is important to be sure that the intervention that is being implemented is improving student outcomes better than expected. By participating in this study, you would be adding to the knowledge in this area and you may receive information that will help you to work more effectively in implementing the HELPS-SG reading intervention with your students. Furthermore, your students who are struggling readers are likely to improve their reading skills.

Confidentiality

The information in the study records will be kept confidential to the full extent allowed by law. Data will be stored securely in a locked room at North Carolina State University that the

principal investigator and faculty sponsor will have access to. No reference will be made in oral or written reports that could link you to the study. The principal investigator will keep a master list of names and ID #s, but she will not share this list with school personnel, and the list will be permanently deleted upon completion of the study. At the beginning of the study you will be assigned a code that will be used to replace your name on all study materials. The master list linking your identity to this code will be kept in a securely locked and stored separate from other study data. This master list will be permanently deleted upon completion of the study.

Compensation

You will not receive financial compensation for participating.

How will participating in this study affect your employment?

Participation in this study is not a requirement of your employment at Wake County Public Schools, and your participation or lack thereof, will not affect your job. No individual reports associated with the data we collect from you will be reported to your principal or other school administrators.

What if you have questions about this study?

If you have questions or concerns at any time about the study or the procedures, you may contact

What if you have questions about your rights as a research participant?

If you feel you have not been treated according to the descriptions in this form, or your rights as a participant in research have been violated during the course of this project, you may contact

Consent To Participate

I have read and understand the above information. I have received a copy of this form. I agree to participate in this study with the understanding that I may choose not to participate or to stop participating at any time without penalty or loss of benefits to which I am otherwise entitled.

Teacher's signature _____

Date _____

Investigator's signature(s) _____

Date _____

Appendix E: Teacher Questionnaire

