ABSTRACT

JOYCE, LAUREN BREEDING. Comparing Two Forms of Feedback to Increase Parents’ Treatment Integrity of a Reading Intervention. (Under the direction of Jeffery Braden and John Begeny).

Performance feedback is an effective procedure for increasing teachers’ treatment integrity of academic and behavioral interventions. Self-evaluation is also a promising strategy for increasing teachers’ treatment integrity of school-based interventions. There is little research, however, examining the effect of performance feedback or self-evaluation on parents’ implementation integrity of home-based interventions. Therefore, the purpose of this study was to compare the effects of performance feedback and self-evaluation on parents’ implementation of a reading intervention to determine if one feedback strategy improved integrity levels more than the other. I used a multiple baseline counterbalanced across-participants design, where parents received weekly performance feedback in one condition and performed weekly self-evaluative feedback in another condition. The results of this comparison were inconclusive, as the majority of the data were either insufficiently stable to compare, or lacked a clear difference between phases. I also predicted that feedback conditions would have higher treatment integrity levels than baseline. Results from four comparisons revealed one reliable difference, showing performance feedback improved integrity over baseline levels. In instances where parents had declining treatment integrity during baseline, treatment integrity improved with the introduction of either feedback method. In general, parents demonstrated improved treatment integrity over time, suggesting that continued feedback supported treatment integrity regardless of feedback type. The results of my study further inform consultation practices for school psychologists working with parents to enhance accurate implementation of home-based reading interventions.
Comparing Two Forms of Feedback to Increase Parents’ Treatment Integrity of a Reading Intervention

by
Lauren Breeding Joyce

A dissertation submitted to the Graduate Faculty of North Carolina State University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Psychology

Raleigh, North Carolina
2015

APPROVED BY:

_____________________________  ________________________________
Jeffery Braden  John Begeny
Committee Co-Chair  Committee Co-Chair

_____________________________
Ann Schulte

_____________________________
Susan Osborne
DEDICATION

My dissertation is dedicated to my children, Holden, Dylan, and Charlotte.
BIOGRAPHY

Lauren Breeding Joyce was born on February 7, 1979 in Olney, Maryland. After spending her early childhood in Silver Spring, MD, Lauren and her family moved to Brookeville, MD. In high school Lauren knew her passion for understanding human behavior and supporting children in need was a natural foundation for a career in psychology. She graduated from Sherwood High School in 1997.

Following high school, Lauren attended The Catholic University of America in Washington, DC where she earned her Bachelors of Arts in Psychology. After earning her undergraduate degree, Lauren attended Pepperdine University in Malibu, California where she earned her Master’s Degree in Clinical Psychology with an emphasis in Marriage and Family Therapy. After working as the Director of Disability Services at Pepperdine University, Lauren moved to Raleigh, North Carolina to earn her PhD in School Psychology at North Carolina State University.

Lauren was married in June 2004 to Jeff Joyce and celebrated 10 years of marriage this past year. She has three children, a son, Holden (age 5), and two daughters, Dylan (age 3) and Charlotte (age 11 months).
ACKNOWLEDGMENTS

My dissertation was completed with the support and encouragement of many people. First, thank you to my advisor, Dr. Jeff Braden. In addition to your time, feedback, and guidance, I am grateful for your contagious enthusiasm. Thank you to my committee, Dr. John Begeny, Dr. Ann Schulte, and Dr. Susan Osborne for your time and feedback throughout my graduate career.

Thank you to my husband, Jeff. Just like Dr. Braden, Jeff encouraged me to “get ‘er done!” I am indebted to Jeff for supporting me throughout graduate school and believing in my abilities. Thank you to my children, who provided the greatest motivation to complete my dissertation and provided the balance I need in my life. Thank you to my mother, father, brothers, mother-in-law, brother-in-law, and sisters-in-law for supporting me throughout my graduate career.

Thank you to my friends and colleagues who believed in me and encouraged me. I owe gratitude to Megan Bennett, Kristen Lewis, and Michelle Tayrose, who continually cheered me on to completion. Your support has been invaluable.

Thank you to Katie Pittman, who took care of my children so I could complete my dissertation. Your dedication to my family provided me the time and peace of mind that allowed me to complete this project.

Finally, thank you to the families who participated in this study. I wouldn’t be finishing my dissertation without you!
# TABLE OF CONTENTS

LIST OF TABLES .......................................................................................................................... xi

LIST OF FIGURES ......................................................................................................................... xiii

Review of the Literature ......................................................................................................................... 1

Consultation ........................................................................................................................................ 1

Overview of Treatment Integrity ........................................................................................................ 2

The Effect of Performance Feedback on Treatment Integrity ......................................................... 4

Schedules of Performance Feedback ................................................................................................. 6

Comparing Performance Feedback with other Feedback Meetings .............................................. 7

Consultee Acceptability of Performance Feedback Meetings ...................................................... 10

Methods of Collecting Treatment Integrity Data ........................................................................... 13

Conclusions on Performance Feedback .......................................................................................... 14

The Effect of Self-Evaluation on Treatment Integrity .................................................................... 15

Parents As Consultees ....................................................................................................................... 18

Statement of the Problem ................................................................................................................ 23

Purpose of Study .............................................................................................................................. 24

Research Questions ......................................................................................................................... 24

Hypotheses ......................................................................................................................................... 25

Method ............................................................................................................................................ 26

Research Design ............................................................................................................................... 26

Participants ....................................................................................................................................... 26
Parent participants ................................................................. 26
Student participants ............................................................... 27
Participant background information ........................................ 27
Inter-group equivalence ........................................................... 29
Setting ..................................................................................... 29
Measures ................................................................................... 30
Treatment integrity ................................................................. 30
Core components ...................................................................... 30
Tips and reminders ................................................................... 31
Intervention session summary .................................................. 31
Treatment acceptability ......................................................... 31
Student performance ............................................................. 32
Comprehensive Test of Phonological Processing (CTOPP) ........ 32
Gray Oral Reading Test, Fifth Edition (GORT-5) ....................... 33
Test of Word Reading Efficiency, Second Edition (TOWRE-2) ... 33
Dynamic Indicators of Basic Early Literacy NEXT, Oral Reading Fluency ......................................................... 34
Procedures ............................................................................... 35
Student participant recruitment ................................................. 35
Parent participant recruitment ................................................. 36
Parent training .......................................................................... 37
Visual analysis of the HELPS tips and reminders treatment integrity.....70

Summary of the visual analysis of tips and reminders.............74

Research questions..................................................................................................................86

Research question 1 ...................................................................................................................86

Research question 2 ...................................................................................................................86

Hypotheses................................................................................................................................87

Hypothesis 1 ..............................................................................................................................87

Hypothesis 2 ..............................................................................................................................88

Hypothesis 3 ..............................................................................................................................88

Hypothesis 4 ..............................................................................................................................89

Supplementary analyses ...........................................................................................................89

Parent participant feedback....................................................................................................91

Parent interview feedback.......................................................................................................95

Student outcomes......................................................................................................................97

Discussion................................................................................................................................100

Student outcomes....................................................................................................................103

Practical implications................................................................................................................103

Limitations...............................................................................................................................106

Future directions and practice................................................................................................108

Conclusion...............................................................................................................................110

REFERENCES........................................................................................................................112
APPENDICES .................................................................................................................. 120

Appendix A: Confidential Screening Questionnaire .............................................. 121

Appendix B: HELPS One-on-One Program: Implementation Protocol .............. 125

Appendix C: Observation Checklist for Completing the HELPS Program .......... 128

Appendix D: Observation Summary Form ................................................................. 130

Appendix E: Observation Summary Form Data....................................................... 131

Appendix F: Intervention Rating Profile-Modified .................................................. 132

Appendix G: Intervention Rating Profile-Performance Feedback ....................... 133

Appendix H: Intervention Rating Profile-Self-Evaluation ..................................... 134

Appendix I: Parent Interview Questions ................................................................. 135

Appendix J: Program Implementation Quiz ............................................................ 137

Appendix K: Scripted Directions .............................................................................. 142

Appendix L: Tips and Reminders for Implementing the Program...................... 144

Protocol ....................................................................................................................... 144

Appendix M: Progress Tracking Form .................................................................... 146

Appendix N: Student Graph .................................................................................... 147

Appendix O: Student's Star Chart ......................................................................... 148

Appendix P: Implementation Flow Chart ............................................................... 149

Appendix Q: Steps and Guidelines for Observing Parents During Their
Implementation of the HELPS Program: Baseline Phase ........................... 151

Appendix R: Steps and Guidelines for Observing Parents During Their
Implementation of the HELPS Program Self-Evaluation

ix
Phase ............................................................................................................. 153

Appendix S: Parent's Observation Checklist for Implementing the
HELPS Program .......................................................................................... 155

Appendix T: Steps and Guidelines for Observing Parents During their
Implementation of the HELPS Program: Performance
Feedback Phase ............................................................................................ 157
LIST OF TABLES

Table 1  Median Integrity Level and Range of Core Components ......................55
Table 2  Case A – Visual Analysis of Core Components ..................................61
Table 3  Case B – Visual Analysis of Core Components ..................................62
Table 4  Case C – Visual Analysis of Core Components ..................................63
Table 5  Case D – Visual Analysis of Core Components ..................................64
Table 6  Case E – Visual Analysis of Core Components ..................................65
Table 7  Case F – Visual Analysis of Core Components ..................................66
Table 8  Case G – Visual Analysis of Core Components ..................................67
Table 9  Case H – Visual Analysis of Core Components ..................................68
Table 10 Case I – Visual Analysis of Core Components ..................................69
Table 11 Median Integrity Level and Range of Tips and Reminders ....................70
Table 12 Case A – Visual Analysis of Tips and Reminders .................................77
Table 13 Case B – Visual Analysis of Tips and Reminders .................................78
Table 14 Case C – Visual Analysis of Tips and Reminders .................................79
Table 15 Case D – Visual Analysis of Tips and Reminders .................................80
Table 16 Case E – Visual Analysis of Tips and Reminders .................................81
Table 17 Case F – Visual Analysis of Tips and Reminders .................................82
Table 18 Case G – Visual Analysis of Tips and Reminders .................................83
Table 19 Case H – Visual Analysis of Tips and Reminders .................................84
Table 20 Case I – Visual Analysis of Tips and Reminders .................................85
Table 21 Parent Participants Acceptability of Performance Feedback ..................92
Table 22  Parent Participants Acceptability of Self-Evaluative Feedback ...............93
Table 23  Parent Participants Acceptability of HELPS Program..........................94
Table 24  Reliable Change Index Score for Students’ Pre- and Post-Achievement Scores..........................................................................................................................97
Table 25  Students Participants’ Pre- and Post-HELPS DORF Scores .........................99
LIST OF FIGURES

Figure 1  Treatment integrity of HELPS Program core components for participants in ABC cohort ..................................................57

Figure 2  Treatment integrity of HELPS Program core components for participants in ACB cohort ..................................................58

Figure 3  Treatment integrity of HELPS Program tips and reminders for participants in ABC cohort ..................................................72

Figure 4  Treatment integrity of HELPS Program tips and reminders for participants in ACB cohort ..................................................73
Review of the Literature

The purpose of the literature review is to introduce the concepts of treatment integrity, performance feedback and self-evaluation. I will begin the literature review by briefly discussing the challenge for school psychologists consulting with parents and how ensuring treatment integrity is critical to determining effects of an intervention. Next I will introduce performance feedback as a method for increasing consultees’ implementation integrity and will review relevant research. The review will include examining the effects of various schedules of performance feedback, consultees’ acceptability of performance feedback, performance feedback compared with other feedback meetings, and methods of collecting treatment integrity data. Following the review on performance feedback, I will discuss self-evaluation as another feedback method for increasing treatment integrity and review relevant research. Finally, I will review research on performance feedback where parents serve as consultees.

Consultation

Although many definitions for consultation exist, the single most defining characteristic of consultation is indirect service delivery (Gutkin & Curtis, 2009). That is, the consultant does not provide direct services to the client, but rather, is typically providing direct support to the consultee, who is, in turn, providing direct services to the client. In school-based consultation, the consultant is typically a school psychologist who is working with a teacher or parent (i.e., the consultee). The teacher or parent then provides direct support to the student, or client. Thus, consultation has two goals. The first is providing services to the client, and the second goal is improving the functioning of the consultee so
that she can prevent or respond more effectively to similar problems in the future (Gutkin & Curtis).

The indirect nature of the consultants’ work has been termed the “paradox of school psychology” (Gutkin & Conoley, 1990, p. 212). Specifically, “to serve children effectively school psychologists must, first and foremost, concentrate their attention and professional expertise on adults” (Gutkin & Conoley, p. 212). The emphasis on supporting the consultee is in part because the consultee (e.g., teacher or parent) greatly influences the child’s environment (e.g., school or home). It is usually the parent or teacher, not the school psychologist, who is implementing the intervention on a day-to-day basis. Through an indirect service delivery model, the school psychologist can exert meaningful change on both the adult in charge of the child’s environment (i.e., consultee) and the child. (Gutkin & Conoley). Therefore, providing an effective intervention for a child typically begins with behavior change on the part of the parent or teacher (Noell et al., 2005).

**Overview of Treatment Integrity**

The focus on caregiver behavior change creates one of the “central challenges confronting school psychologists treating children” (Noell et al., 2005, p. 88). That is, the school psychologists’ challenge lies in trying to ensure that the intervention is implemented as intended—ensuring treatment integrity. Treatment integrity has been defined as the degree to which an intervention is implemented as intended (Gresham, Gansle, & Noell, 1993; Moncher & Prinz, 1991). Recently, however, a new working definition of treatment integrity was proposed, introduced to better capture the multidimensional conceptualization of this construct (Sanetti & Kratochwill, 2009). In this proposed working definition, treatment
integrity is defined as “the extent to which essential intervention components are delivered in a comprehensive and consistent manner by an interventionist trained to deliver the intervention” (Sanetti & Kratochwill, p. 448).

The earliest definitions of treatment integrity conceptualized it as a single construct—adherence to the intervention components (e.g., Peterson, Homer, & Wonderlich, 1982). However, several researchers have expanded on the conceptualization of treatment integrity by defining it as a multidimensional construct (e.g., Dane & Schneider, 1998; Waltz, Addis, Koerner, & Jacobson, 1993). For example, in addition to treatment adherence, Waltz and colleagues identified therapist competence as another dimension of integrity that should be measured. Dane and Schneider (1998) identified five dimensions of integrity, including adherence, exposure, quality of delivery, responsiveness, and program differentiation. Despite the identification of these additional aspects of integrity, adherence remains the most commonly measured dimension of integrity in research studies (Schulte, Easton, & Parker, 2009).

Treatment integrity has important research and practice implications. From a research perspective, treatment integrity influences the internal and external validity of a research study (Gresham et al., 1993; Moncher & Prinz, 1991; Sanetti & Kratochwill, 2005). Internal validity is compromised if integrity of implementation is not monitored, as the accuracy of independent variable implementation is unknown. It is then difficult to attribute the change in the dependent variable to the manipulated changes in the independent variable (Gresham et al.). Moncher and Prinz agree, and add that the absence of treatment integrity monitoring prohibits the examination of possible “contaminants added to the treatment” (p. 249).
External validity is also compromised if implementation integrity is not monitored, as interventions cannot be replicated with precision (Noell, Gresham, & Gansle, 2002). Failure to monitor treatment integrity “compromises the science of building a replicative history” (Gresham et al., 1993, p. 258). If treatment integrity has not been monitored for a specific study, consumers of the study will not know what was done nor how well it was done, and will therefore be unable to understand whether outcomes are linked to interventions (internal validity), and to generalize the findings of the research (external validity).

From a practice perspective, monitoring treatment integrity is critical for documenting response to intervention. A student’s response to intervention determines his level of intervention and movement within the tiers of the problem-solving model. If the intervention is not implemented with high fidelity it is difficult to confidently attribute the child’s lack of or decreased response to the intervention as failure to acquire the skills (Roach & Elliott, 2008).

**The Effect of Performance Feedback on Treatment Integrity**

The importance of increasing levels of treatment integrity and maintaining high levels of treatment integrity has led to research examining how consultants influence treatment integrity of intervention implementation. One method of supporting the consultees’ implementation of the intervention plan is through performance feedback. Performance feedback has been defined as “monitoring a behavior that is the focus of concern and providing feedback to the individual regarding that behavior” (Noell et al., 2005, p. 88). Performance feedback usually includes (a) data review, (b) praise for accurate implementation, (c) corrective feedback, and (d) response to questions and comments.
At least within the school-based consultation research, providing performance feedback to the consultee increases treatment integrity, with the presentation of treatment fidelity data serving as the “key element” (Sanetti, Luiselli, & Handler, 2007, p. 456).

A series of research studies were conducted with a purpose of better understanding how consultation follow-up processes influence implementation integrity (Codding et al., 2005; Codding, Livanis, Pace, & Vaca, 2008; DiGennaro, Martens, & Kleinmann, 2007; DiGennaro, Martens, & McIntyre, 2005; Jones, Wickstrom, & Friman, 1997; Mortenson & Witt, 1998; Noell, Duhon, Gatti, & Connell, 2002; Noell, Witt, Gilbertson, Ranier, & Freeland, 1997; Witt, Noell, LaFleur, & Mortenson, 1997; Noell, Witt, LaFleur, Mortenson, Ranier et al, 2000; Noell et al., 2005). In each of these studies, researchers examined the influence of performance feedback on consultee implementation integrity, with classroom teachers serving as the consultees in all of the studies. Across all studies, after a decline in treatment integrity during the baseline phase, the introduction of performance feedback resulted in increased treatment integrity. Performance feedback led to increases in treatment integrity when consultants provided it vocally (Codding et al., 2008; Codding et al., 2005; Jones et al, 1997; Mortenson & Witt, 1998), and in a combination that included treatment integrity data presented in a graph (DiGennaro et al., 2007; DiGennaro et al. 2005; Noell et al., 2005; Noell et al., 2002; Noell et al., 2000; Noell et al., 1997; Witt et al., 1997). In all but two studies (i.e., Codding et al., 2008; Codding et al., 2005) performance feedback meetings included the consultant presenting consultee implementation data and student performance data to the consultee.
Schedules of Performance Feedback

In addition to examining the overall effect of performance feedback on implementation integrity, researchers in several studies specifically examined whether modifying the schedules for providing performance feedback influenced treatment integrity (e.g., Coddington et al., 2005; DiGennaro et al., 2005; DiGennaro et al., 2007; Mortenson & Witt, 1998; Noell et al., 2000). Determining that leaner schedules of performance feedback result in similar treatment integrity levels that followed daily performance feedback is important, as consultation practices can be more efficient (Mortenson & Witt). For example, school psychologists working in more than one school or working in rural settings where geographical distances prevent multiple visits per week need consultation strategies that are effective (Mortenson & Witt). In the majority of studies, researchers progressively reduced, or thinned, the schedule for providing performance feedback from daily meetings to once every two days, once every three days, once a week, and/or once every two weeks. When the original schedules of daily performance feedback meetings were thinned, most consultees maintained high levels of treatment integrity (DiGennaro et al., 2007; DiGennaro et al., 2005; Noell et al., 2002; Noell et al., 2000).

Rather than progressively thinning the schedule of performance feedback, Mortenson and Witt (1998) provided consultees with weekly performance feedback from the outset of the performance feedback phase. Although they found that the introduction of performance feedback resulted in immediate increases in treatment integrity levels, the effect was not as large as the findings reported in an earlier study that examined daily performance feedback (Witt et al., 1997). Coddington and colleagues (2005) extended the research on schedules of
performance feedback by evaluating the timing effects of performance feedback (i.e., when it was provided to the teacher the same day as the direct observation, but was only observed and provided every other week). Following a decline in baseline integrity level, the introduction of performance feedback resulted in improved teacher intervention implementation. Compared to the results found in a previous study (Mortenson & Witt, 1998), it appears that teachers in the Codding and colleagues (2005) study implemented the intervention with overall greater integrity and with less variability. Thus, these results provide some initial evidence for the efficacy of performance feedback on treatment implementation when provided the same day as the direct observation, every other week.

There are inconsistent results across the treatment integrity data in relation to the schedule of performance feedback, especially regarding the effectiveness of differing schedules. Furthermore, additional research is warranted that examines performance feedback when it is provided the same day as the direct observation. Research on these issues may help consultants provide more effective and efficient feedback (Noell et al., 2000), as less frequent performance feedback may be as (or more) effective as providing feedback daily or every other day.

Comparing Performance Feedback with other Feedback Meetings

After initial studies provided results that supported the efficacy of performance feedback, subsequent studies (DiGennaro, et al., 2007; Noell, et al., 2005; Noell et al., 2002; Noell et al., 2000) compared performance feedback with other types of feedback meetings. The comparative feedback meetings included conducting brief follow-up meetings with the consultees, where the consultant inquired about how the intervention was going and whether
the consultee had any questions (Noell et al., 2000), or reviewing a permanent product to
determine intervention implementation based the consultees’ completion of the permanent
product (Noell et al., 2002). More recently, performance feedback has been compared with
two other feedback conditions in the same study (Noell et al., 2005) and a condition that
focused on student performance (DiGennaro et al., 2007). What follows is a discussion of the
Noell et al. (2002) and DiGennaro et al. (2007) studies in greater detail.

As the first randomized design study that examined performance feedback with other
conditions, Noell and colleagues (2005) compared three different feedback conditions:
weekly follow-up, commitment emphasis, or performance feedback. Conceptualized as the
“best practice approach” for consultation in the schools, participants in the weekly follow-up
condition met briefly with the consultant, who inquired about to the extent of intervention
implementation, the improvement of the student, and whether the teacher had any questions
or concerns (Noell et al., p. 95). Participants in the commitment emphasis condition met
weekly with the consultant and covered the same components as those covered in the weekly
follow-up condition but consultants also included a social influence procedure designed to
strengthen the link between teachers’ commitment to implement the intervention and their
actual implementation. The consultant did not share any implementation integrity data with
the teachers in either the weekly follow-up or commitment emphasis conditions. Finally,
participants in the performance feedback condition met with the consultant who reviewed the
intervention permanent products and graphed student behavior and intervention
implementation. The consultant also praised the consultee for correct implementation of the
intervention steps and identified which steps were missed or implemented incorrectly.
Performance feedback meetings were thinned from daily to every other day to weekly based on the teachers’ treatment integrity levels.

Intervention implementation in the performance feedback condition differed significantly from intervention implementation in both the weekly follow-up and commitment emphasis conditions. An ANOVA revealed a significant main effect for condition, indicating that treatment integrity levels in the performance feedback condition were substantially higher than levels in the other two conditions [F(2, 42) = 9.0, p = .001] (Noell et al., 2005). A limitation to this study, however, is that participants in the performance feedback group had more contact with the consultant from the outset of the study than did participants in the other two conditions; thus, the strength of performance feedback when it is provided weekly from the outset is unknown.

A more recent study (DiGennaro et al., 2007) compared performance feedback to a condition in which teachers set a student behavioral goal and only received feedback on student behavior. The performance feedback condition, labeled teacher performance feedback and directed rehearsal with meeting cancellation (PF + DR/MC), had multiple components. First, consultants provided daily written feedback on teacher implementation and student performance. If integrity was not 100%, teachers met with the consultant the following day to review and practice the missed/incorrect steps (i.e., directed rehearsal). In the goal setting and student performance feedback condition (Goal Setting + PF) teachers set a student behavioral goal and received daily written feedback only on student performance. Teachers implemented the intervention with higher integrity during the PF + DR/MC phase. Although teachers demonstrated higher implementation levels during the Goal Setting + PF
phase than during baseline implementation levels, the implementation levels during the PF +
DR/MC phase remained higher than those during the Goal Setting + PF phase. Furthermore,
the researchers demonstrated within-subject functional control with two teachers by
conducting a brief reversal to the PF + DR/MC phase following the Goal Setting + PF phase.
When compared with other feedback meetings, performance feedback resulted in higher
levels of treatment integrity. A commonality across the studies is the absence of reviewing
treatment integrity data during the comparative feedback meetings and the review of
treatment integrity data and use of graphs to present treatment integrity data during the
performance feedback meetings. It appears that presentation of treatment integrity data is
necessary for improving intervention implementation; however, whether it is necessary that
data is presented in a graph format or not is unclear, as several researchers found an increase
in treatment integrity following performance feedback meetings that did not include the
graphic presentation of treatment integrity (e.g., Codding et al., 2008; Codding et al., 2005;
Jones et al., 1997; Mortenson & Witt, 1998). As performance feedback is a multi-component
feedback method, it is wise to further examine the individual effects of each component to
determine which of the components are critical to influencing behavior change.

Consultee Acceptability of Performance Feedback Meetings

More than a decade ago, Mortenson and Witt (1998) speculated as to how
performance feedback meetings functioned (i.e., as a positive reinforcer, a negative
reinforcer, a prompt). They discussed the possibility that performance feedback
contingencies may be negatively reinforcing. That is, teachers may experience the
performance feedback meeting as aversive and, thus, teachers may increase the accuracy of
intervention implementation to avoid future meetings with the consultant. To determine whether performance feedback contingencies are negatively reinforcing, Mortenson and Witt recommended that future studies of performance feedback also include teacher assessments of the acceptability of performance feedback.

Overall, consultees rate performance feedback procedures favorably (Codding et al., 2008; Codding et al., 2005; DiGennaro et al., 2007; DiGennaro et al., 2005; Noell et al., 2005; Noell et al., 2002). In the majority of studies that assessed consultee opinions on performance feedback procedures (Codding et al., 2008; Codding et al., 2005; Noell et al., 2005; Noell et al., 2002) consultees were required to meet with the consultant. In two studies, however, researchers used a negative reinforcement contingency with their participants (DiGennaro, et al., 2007; DiGennaro, et al., 2005). That is, teachers could avoid meeting with the consultant if their implementation accuracy remained at 100%. Consultees received daily written feedback on their implementation integrity and only had direct contact with the consultant if their integrity level dropped below 100%. When consultees met with the consultant, consultees were required to practice missed steps of the intervention three times (i.e., directed rehearsal). Directed rehearsal was a unique component to the performance feedback procedures in the DiGennaro et al. studies (2007; 2005).

Consultee ratings of the performance feedback procedures that used a meeting cancellation contingency were positive, as all teachers agreed or strongly agreed with the items pertaining to the perceived benefit of the consultants’ procedures and liking the procedures used by the consultant to help them implement the intervention (DiGennaro et al., 2007; DiGennaro et al., 2005). All consultees agreed that receiving daily written feedback
was an acceptable way to provide treatment integrity data. Despite overall positive ratings, the DiGennaro et al (2007) study reported some specific negative ratings of performance feedback. Some participants disagreed with practicing missed steps of the intervention and avoiding a meeting with the consultant (both components of the performance feedback condition). Furthermore, when compared with another feedback condition where consultees received daily written feedback on student performance and were not required to meet with the consultant, consultees rated the performance feedback condition slightly less favorably than the student performance feedback condition (DiGennaro et al., 2007).

It is unclear whether avoiding performance feedback meetings contingent on treatment integrity levels is negatively reinforcing. On one hand, consultees who were able to avoid meeting with the consultant based on high treatment integrity levels rated these procedures favorably. Furthermore, when compared to another feedback condition, consultees had higher treatment integrity levels during the condition in which they could avoid meeting with the consultant (DiGennaro et al., 2007). On the other hand, consultees who met with consultants for performance feedback meetings also rated performance feedback procedures favorably. It is difficult, therefore, to attribute the higher levels of integrity during the performance feedback contingency condition directly to the meeting cancellation variable because the researchers did not isolate this one variable. For example, providing the consultant with daily teacher implementation data, not the aversion to meeting with the consultant, may have been the active component in eliciting changes in teacher intervention implementation. Or, it may be that performance feedback when combined with directed rehearsal is perceived negatively, and thus, consultees will ensure high treatment
integrity to avoid practicing missed intervention steps. DiGennaro and colleagues (2007) noted that future research should include the examination of multicomponent consultant procedures in isolation before they are combined into a packaged procedure.

**Methods of Collecting Treatment Integrity Data**

With the exception of two studies (DiGennaro et al., 2005; Jones et al., 1997), researchers collected treatment integrity data using permanent products. The major benefit of permanent products over direct observation is reduced observer reactivity. Observer reactivity occurs when the act of observation changes the behavior being observed; for example, when an interventionist implements the intervention with higher integrity only when someone watches (Gresham et al., 2000). In the two studies that used observation to collect treatment integrity data, it is possible that teachers executed the intervention more precisely during these observations, but had more variable implementation during the sessions in which they were not observed.

One study (Codding et al., 2008), however, specifically examined the role of observer reactivity when conducting a direct evaluation of intervention implementation. During the course of this study the observer collected half of the observation data inside the classroom, where the teacher was aware of the observer’s presence, and half of the observation data outside of the classroom (but observing through a one-way mirror). The days, times, and conditions (observer-present vs. absent) for observations were randomly selected to reduce the likelihood that teachers would become aware of when/how the observations would take place. There was no difference in treatment integrity between the two conditions (observer-present & observer-absent) during both the baseline and the performance feedback phases.
Thus, the data suggest that observer reactivity in response to direct observation does not lead to positive effects of performance feedback (Codding et al.).

Although the use of permanent products may offer the advantage of reducing observer reactivity, this method does have one substantial drawback. Permanent products may fail to capture many elements of an intervention. For example, a worksheet completed by student during an intervention session shows that the intervention was delivered to the student but it does not reflect how the intervention was delivered. That is, the worksheet may not capture interventionist behaviors that are indicative of accurate implementation of the intervention. Therefore, it is recommended that researchers use several methods for monitoring treatment integrity (Gresham et al., 2000; Moncher & Prinz, 1991), with the data from permanent products used in conjunction with direct measures of implementation (e.g., observation) (Gresham et al.).

**Conclusions on Performance Feedback**

Performance feedback is associated with increased levels of intervention implementation when compared to baseline implementation levels (Codding et al., 2008; Codding et al., 2005; DiGennaro et al., 2007; DiGennaro et al., 2005; Mortenson & Witt, 1998; Noell et al., 2005; Noell et al., 2002; Noell et al., 2000; Noell et al., 1997; Witt et al., 1997). Performance feedback is associated with higher levels of implementation for several types of interventions, including behavior-management interventions (Codding et al., 2005; Codding et al., 2008; DiGennaro et al., 2007; DiGennaro et al., 2005; Noell et al., 2002), academic interventions (Mortenson & Witt, 1998; Noell et al., 2000; Witt et al., 1997), the use of contingent praise (Jones et al., 1997) and peer tutoring (Noell et al., 2000).
Furthermore, when compared with other feedback conditions, performance feedback is associated with higher levels of implementation integrity (DiGennaro et al., 2007; Noell et al., 2005; Noell et al., 2002; Noell et al., 2000). Finally, consultees rate performance feedback procedures favorably (Codding et al., 2008; Codding et al., 2005; DiGennaro et al., 2007; DiGennaro et al., 2005; Noell et al., 2005; Noell et al., 2002), suggesting that consultees find the procedures effective and experienced meetings with the consultant positively.

**The Effect of Self-Evaluative on Treatment Integrity**

Self-evaluation is another method in which an interventionist can receive feedback on her implementation of an intervention. Self-evaluation and self-monitoring have been used somewhat synonymously in the literature; however, much of the research on self-monitoring involves real-time recording of one’s own behavior (e.g., Kalis, Vannest, & Parker, 2007; Richman, Riordan, Reiss, Pyles, & Bailey, 1988). The focus on self-evaluation in this document refers to an interventionist reviewing and recording the occurrence/nonoccurrence of her behavior of interest from an audio or videotape.

From the consultant’s perspective, self-evaluation may be a more time efficient method for ensuring treatment integrity because the consultant does not have to be present to observe intervention implementation and provide corresponding feedback. However, there are few published studies that examine the effect of self-evaluation on interventionists’ treatment integrity of school-based interventions. That is, does evaluating oneself executing the steps of an intervention via audio or videotape increase treatment integrity (provided that integrity is not already 100%)?
Researchers have examined the effect of self-evaluation on the implementation integrity of teachers’ use of praise (Sutherland & Wehby, 2001), Discrete Trial Instruction (DTI) (Belfiore, Fritts, & Herman, 2008), and a behavioral intervention (Pelletier et al., 2010). Both Belfiore and colleagues and Pelletier and colleagues used video self-monitoring, whereas Sutherland and Wehby used audio recordings. In all three studies teachers reviewed the audio or videotape of their intervention to determine the level of intervention implementation. The frequency that participants reviewed their audio/videotape differed across studies, as participants either reviewed audio/videotapes of themselves daily (Belfiore et al.; Sutherland & Wehby) or only once or twice during the intervention phase (Pelletier et al.).

In all three studies (Belfiore et al., 2008; Pelletier et al., 2010; Sutherland & Wehby, 2001) researchers compared the participants’ performance during a baseline phase to performance during the intervention phase. In two studies (Belfiore et al.; Sutherland & Wehby) experimenters did not provide any feedback to participants who were self-evaluating. In one study (Pelletier et al., 2010), participants watched and scored one of their own videos alongside one of the experimenters. If the participant and the experimenter disagreed on the scoring of procedural integrity, the experimenter explained to the participant why she scored the measure differently. During the self-evaluation phase, participants in the Pelletier et al. study only viewed baseline videotapes of themselves, whereas participants in the remaining studies reviewed audio/videotapes of themselves from the same day that they were audio/videorecorded.
Findings from these studies suggest that self-evaluation increases treatment integrity (Belfiore et al., 2008; Pelletier et al., 2010; Sutherland & Wehby, 2001). Across these studies, participants increased their treatment integrity after the introduction of self-evaluative feedback. For participants in the Sutherland and Wehby study, that used a between groups design, there was a reliable increase in total praise (the interventionist behavior of interest) from pretreatment to treatment that favored the self-evaluation group \((F(1,18) = 23.42, \ p < .001, \ ES = 2.0)\). All participants in the remaining two studies (Belfiore et al., 2008; Pelletier et al., 2010) increased their integrity level above baseline levels after the introduction of self-evaluation.

One major limitation to the study by Pelletier and colleagues (2010) was the combination of self-evaluation and direct feedback in the performance feedback phase. The use of a multicomponent intervention package makes it difficult to determine which components are the active agents responsible for change. Furthermore, researchers in this study did not collect data on the accuracy with which participants scored the videos in the pre-training and self-evaluation phases, and therefore, how much feedback participants received. Thus, the positive results from this study cannot be directly attributed to self-evaluation alone but may be a combination of self-evaluation and experimenters’ vocal feedback.

In one study researchers also assessed treatment acceptability (Sutherland & Wehby, 2001). Participants completed the Treatment Acceptability Rating Form-Revised where they rated the acceptability and feasibility of the self-evaluation intervention. Scores on this measure suggested that teachers would be willing to implement the self-evaluation
intervention again, and expected the intervention to be effective for their students. Furthermore, teachers perceived minimal adverse consequences to using self-evaluation. Thus, teachers rated the acceptability of the intervention favorably.

Overall, self-evaluation is a promising method for interventionists to gain feedback on their intervention implementation. The introduction of self-evaluation resulted in increased treatment integrity when compared to baseline levels (Belfiore et al., 2008; Pelletier et al., 2010) and resulted in a reliable difference in treatment integrity between participants who used self-evaluation and participants who did not (Sutherland & Wehby, 2001). Furthermore, participants rated the self-evaluation intervention favorably (Sutherland & Wehby). The preliminary evidence for the efficacy of self-evaluation provides a base for additional research, specifically parents’ use of self-evaluation.

Parents as Consultees

In addition to the school psychologist’s role of working with teachers and school staff, school psychologists also enhance student outcomes by cultivating home-school partnerships via parent consultation. Encouraging parents to be involved in their child’s education is critical to educational success (Buerkle, Whitehouse, Christenson, 2009), as parental involvement in subject matter areas (e.g., reading) has been linked to positive outcomes for the student (Resetar, Noell, & Pellegrin, 2006). As discussed previously, the indirect nature of consultation can present challenges for school psychologists as they seek to ensure the integrity of home-based interventions. Although research findings support the use of performance feedback and self-evaluation with teachers to increase the treatment integrity...
of interventions, the influence of performance feedback or self-evaluation on parents’ implementation of interventions is largely unknown.

Of the studies that have examined the effect of performance feedback on caregivers’ treatment integrity, the data are not reflective of the findings where teachers serve as the interventionist. For example, Connell (2010) examined the effects of performance feedback on a grandmother’s implementation integrity of a home-based behavior intervention. The clients were the grandmother’s two grandchildren who resided with her. The consultant provided weekly performance feedback that consisted of reviewing the permanent products, graphing the consultee and clients’ behavior, problem-solving implementation and data collection barriers, praising correctly implemented intervention steps, and identifying intervention steps that were omitted or implemented incorrectly. When implementation integrity fell to 0% the consultant initiated weekly follow-up phone calls to the consultee. During the phone calls the consultant inquired as to whether the data collection sheets were being filled out daily, provided positive feedback if the sheets were being filled out daily, and provided recommendations to increase implementation if sheets were not completed daily.

Overall, the data from this study are not consistent with data from previous studies on performance feedback. Although treatment integrity declined during the baseline phase, the daily integrity level was never above 50% from the outset of the study. On the tenth day of the study, integrity level fell to 0% and remained at that level for about thirty days. The introduction of performance feedback resulted in an immediate increase in treatment integrity (to 100%) but it decreased immediately to 0% thereafter. A second performance feedback meeting did not influence integrity level, as it remained at 0%. The consultant’s initiation of
phone call follow-up meetings resulted in variable data, where integrity levels fluctuated between either 0% or 100%. Thus, integrity data reflects an all-or-nothing pattern, where it appears that intervention implementation occurred in its entirety or did not occur at all. It is interesting, however, that data for both clients show that points were earned on days in which the consultee reported 0% implementation of the intervention. The inconsistency between the client data and the treatment integrity data may reflect poor record-keeping on the part of the grandmother. The sole reliance on permanent products to collect treatment integrity data, however, prevents further investigation into the accuracy with which the grandmother actually implemented the intervention.

Despite results that demonstrated that treatment integrity was higher, albeit highly variable, with the use of performance feedback and phone call follow-ups than without, there are several limitations to this study (Connell, 2010). First, it is unclear whether the consultant provided performance feedback meetings during the phase that also included phone call follow-up meetings, as the performance feedback meetings are not indicated in the graphed data whereas the phone calls are. Second, the study lacks strong internal validity due to the inconsistency between treatment integrity data and client behavior data (i.e., points earned by clients). Because the only measure of treatment integrity is based on permanent products, there is no way to reconcile the inconsistencies in the data. Although a behavioral intervention may be challenging to directly observe because the consultee monitors the clients’ behavior all day and does not provide points for appropriate behavior until the end of the day, the consultant could have observed the consultee completing the first part of the intervention (i.e., reminding the clients of the house rules each morning) to determine that
this component could be implemented correctly. Finally, the external validity of this study is compromised due to the methodological design. That is, Connell (2010) did not replicate the findings of this study to rule out timing confounds, as would be done in a multiple baseline design.

A study by LeBlanc (2009) examined the effects of performance feedback on parents’ implementation of a reading fluency intervention. Performance feedback and performance feedback plus a graphing component were compared to determine if one was superior to the other in increasing treatment integrity level. After being trained in the intervention, parents implemented the intervention during the baseline phase, and the consultant provided support to all of the parent participants, either in person or via phone, during the this phase. The consultant provided performance feedback after the implementation level fell below 80% or demonstrated a decreasing trend during the baseline phase. Performance feedback consisted of providing the parent her implementation score for each session. When the integrity level declined or consistently fell below 80% during the performance feedback phase, the consultant provided parents with performance feedback plus a graph of the parents’ integrity score.

The data from this study provide inconclusive results. Across participants, the use of performance feedback did not consistently result in significantly higher levels of treatment integrity, as the treatment integrity levels of each participant differed. One participant had an initial high treatment integrity level during the baseline phase that declined by the second session. Providing performance feedback and performance feedback plus graphing did not have any discernible effect. The data for another participant demonstrated a decreasing trend
for treatment integrity during the performance feedback phase. When the graphing
component was added, treatment integrity increased to 100%. However, after one
performance feedback plus graphing session, another adult in the home took over
implementing the intervention with the child and the remaining data could not be included in
the study. The third participant demonstrated increased treatment integrity during the
performance feedback phase (80% and above) but discontinued implementing the
intervention after the seventh feedback meeting due to her child’s behavior problems. Thus,
the participant did not receive the graphing component. The fourth participant demonstrated
wide variability in her integrity levels across all phases, leading to inconclusive results. The
final participant demonstrated a gradual increasing trend of integrity during the baseline
phase followed by an initial increase, albeit insignificant, following the introduction of
performance feedback. Integrity, however, gradually declined and the introduction of the
graphing component resulted in an increase in integrity, although integrity level did not
surpass the 80% criterion level.

Several conditions limit the interpretability of the results. First, two of the five parents
reported that their child’s problem behaviors affected the implementation of the intervention.
Child behavior (e.g., compliance, attentiveness) during implementation of the intervention
was not assessed however. Therefore, data on these two parents must be interpreted with
cautions because another variable (child behavior) may have been more influential than
performance feedback in affecting implementation integrity. Second, participation issues
with three of the parents led to inconclusive data. Two parents were inconsistent in attending
meetings and maintaining communication with the consultant. Another parent allowed
another adult in the home to implement the intervention with the child while she discontinued
her participation, thus invalidating the integrity data from that point on. Third, some of the
parents were provided performance feedback in person while others were provided feedback
over the phone. The mode in which performance feedback was provided may have
influenced its effectiveness. Fourth, performance feedback was intended to be provided
weekly. Due to scheduling challenges and other issues related to parent participation,
however, performance feedback did not follow this schedule for all parents.

The findings from the two aforementioned studies (Connell, 2010; LeBlanc, 2009)
provide inconclusive results about the effectiveness of performance feedback on increasing
parents’ implementation integrity. Major limitations in both studies preclude solid findings
and diminish the external validity of these studies. Given the importance of parent
involvement in their child’s education, and the importance of ensuring that parents
implement home-based academic interventions with high integrity, the issue of feedback
effects on parents’ treatment integrity is a worthy research topic. Therefore, it is critical that
additional research is conducted that examines parents’ fidelity of intervention
implementation following performance feedback.

**Statement of the Problem**

Research findings support the use of performance feedback as an effective means for
increasing teachers’ treatment integrity of school-based interventions. Little research,
however, has been conducted that examines the effect of performance feedback on parents’
implementation of an intervention, especially when they implement interventions in the home
setting. Additionally, very few studies have isolated and compared different components of
performance feedback. In many studies, performance feedback is provided by consultants to consultees (e.g., teachers or caregivers) as a multi-component package that may include vocal feedback, data on the consultees’ implementation integrity, data on student performance, graphs displaying integrity and student performance data, identification of the missed intervention steps, discussing methods for enhancing implementation of missed steps, and practicing missed steps of the intervention. Thus, it is difficult to discern which of these components may be the most effective or if other feedback mechanisms, such as self-evaluation, are equally effective.

**Purpose of the Study**

The purpose of this study was to compare the effects of two types of performance feedback on treatment integrity. Specifically, I examined whether parents demonstrated higher treatment integrity levels when provided weekly vocally-mediated 3rd party performance feedback (standard performance feedback; hereafter referred to as performance feedback) or when provided weekly vocally-unmediated feedback (i.e., feedback via self-evaluation; hereafter referred to as self-evaluation). A secondary purpose of the study was to compare treatment integrity with the absence of feedback to treatment integrity when provided with either performance feedback or self-evaluation. Lastly, I evaluated the effects of the Helping Early Literacy with Practice Strategies (HELPs) Program intervention on students’ oral reading fluency and comprehension.

**Research Questions**

This study was conducted to answer the following research questions and hypotheses:
1. Is treatment integrity of the HELPS intervention core components higher when parents receive performance feedback than when parents receive self-evaluative feedback?

2. Is treatment integrity of the HELPS intervention tips and reminders higher when parents receive performance feedback than when parents receive self-evaluative feedback?

**Hypotheses**

1. Treatment integrity of the HELPS intervention core components will be higher in the performance feedback condition than in the baseline condition.

2. Treatment integrity of the HELPS intervention tips and reminders will be higher in the performance feedback condition than in the baseline condition.

3. Treatment integrity of the HELPS intervention core components will be higher in the self-evaluation feedback condition than in the baseline condition.

4. Treatment integrity of the HELPS intervention tips and reminders will be higher in the self-evaluation feedback condition than in the baseline condition.
Method

Research Design

I used a multiple baseline across subjects design. My study included three phases: (a) Implementation Baseline, (b) Self-Evaluative Feedback, and (c) Performance Feedback. The latter two phases were counterbalanced across the participant dyads so that five of the parent participants proceeded through the phases in the following order: ABC. The remaining four parent participants proceeded through the phases in the following order: ACB. I randomly assigned parent participants to one of the two groups by matching student participants into dyads based on fall benchmark oral reading fluency scores. I flipped a coin to determine group assignment, whereby the first parent/child dyad was assigned to either group based on flipping a heads (ABC phase sequence) or a tails (ACB phase sequence) and the second dyad was placed in the other group. I staggered the introduction of the experimental phases, so that participants in the same group (i.e., ABC or ACB) did not begin an experimental phase on the same session number.

Participants

Parent participants. Nine parents participated in this study. Prior to the study no parent had implemented the HELPS reading program with their child. All parents completed a self-report questionnaire (see Appendix A), similar to the questionnaire used by Mitchell and Begeny (2014). The questionnaire sought information on parent participants’ ethnicity, gender, marital status, level of formal schooling, occupation, and reading activities conducted in the home. The questionnaire also sought information on student participants’ demographic information, medical and educational history, and reading activities in the home.
Parent participants included eight females and one male. The mean age of parent participants was 39.7 years ($SD = 2.5$). Two parents were African American and seven parents were Caucasian. Parents’ average length of schooling was 15.6 years ($SD = 1.58$; range = 14-19 years). All of the parent participants were the biological parent of the student participant.

**Student participants.** Nine students participated in this study. Student participants were recruited from an elementary school in a rural county of North Carolina. Student participants were $2^{\text{nd}} (N=5)$, $3^{\text{rd}} (N=3)$, and $4^{\text{th}} (N=1)$ grade students who were at risk or struggling readers (i.e., those at the $10^{\text{th}}$ - $50^{\text{th}}$ percentile on a universal screening measure of oral reading fluency). Six student participants were female and three were male. The mean age of student participants was 7.7 years ($SD = .97$). Two students were African American, six students were Caucasian, and 1 student was multiracial (Asian/Caucasian).

**Participant background information.** Parents were asked to indicate the frequency with which they engaged in the following activities related to reading. Using a 4-point Likert scale (4 = almost daily, 3 = once or twice a week, 2 = once or twice a month, 1 = hardly ever) parents indicated the following about home activities: their child saw an adult read a book/newspaper/magazine typically once or twice a week ($M = 3.22; SD = 1.09$; range = 1-4); their child saw an adult read on the computer about once or twice a week ($M = 3.55; SD = 1.01$; range = 1-4); their child asked for help with homework once or twice a week to almost daily ($M = 3.75; SD = 0.46$; range = 3-4); parents monitored their child’s homework almost daily ($M = 4.0; SD = 0.0$; range = 4-4); parents read aloud to their child typically once or twice a week ($M = 3.33; SD = 1.00$; range = 1-4); their child read a book aloud to them
typically once or twice a week ($M = 3.66; SD = 0.50; range = 3-4$); and their child’s teacher asked them to read with their child or had homework assignments that required parent involvement typically once or twice a week to almost daily ($M = 3.77; SD = 0.66; range = 2-4$).

When parents were asked about activities that their child may do parents indicated their child: visited the community library and brought home a book hardly ever ($M = 1.33; SD = 0.71; range = 1-3$); typically read signs, labels, and grocery lists once or twice a week ($M = 3.11; SD = 0.78; range = 2-4$); played educational games hardly ever ($M = 1.77; SD = 0.44; range = 1-2$); played computer games involving reading typically once or twice a month ($M = 2.44; SD = 1.01; range = 1-4$); wrote a story/note/card/diary once or twice a month ($M = 2.44, SD = 1.01; range = 1-4$); read recipe/craft/game instructions with their parent typically hardly ever ($M = 1.87; SD = 0.64; range = 1-3$); read a book for fun typically once or twice a month ($M = 2.22; SD = 0.83; range = 1-3$); watched educational TV shows with an adult typically once or twice a month ($M = 2.11; SD = 0.93; range = 1-4$).

Finally, parents were asked to rate how frequently they used different reading activities with their children. Parents indicated that they: taught words on flashcards hardly ever ($M = 1.67; SD = 1.00; range = 1-3$); asked their child to read aloud for practice typically once or twice a week ($M = 3.67; SD = 0.50; range = 3-4$); assisted with new word by sounding out letters typically once or twice a week to almost daily ($M = 3.78; SD = 0.44; range = 3-4$); assisted with a new word by using the meaning of the sentence typically once or twice a month ($M = 2.44; SD = 1.01; range = 1-4$); assisted with a new word by asking their child to try the word again once or twice a week to almost daily ($M = 3.78; SD = 0.44$);
range = 3-4); assisted with new word by providing the correct word typically once or twice a week ($M = 3.56; SD = 0.73; range = 2-4)$.

**Inter-group equivalence.** I conducted preliminary analyses on demographic variables of both the parent participants and the student participants to assess whether there were significant differences between the dyads that received the self-evaluative feedback first and the dyads that received the performance feedback first. I conducted a $t$-test on each of the items that assessed two variables: parent reading activities in the home and student reading activities in the home. All $t$-tests assessing parent reading activities in the home showed no reliable differences at the 0.05 probability level with 1 degree of freedom, $t = -0.88 - 2.45$, $NS$. All $t$-tests assessing student reading activities in the home showed no reliable differences at the 0.05 probability level with 1 degree of freedom, $t = -0.08 - 1.63$, $NS$. I conducted chi-square analyses ($x^2$) to examine parent education level, student gender, and student ethnicity. For student gender, I decided that there was a significant difference between the groups if the calculated chi-square ($x^2$) statistic exceeded the critical value for the 0.05 probability level (3.84). The difference between the groups was not significant $x^2 (1, N = 9) = 0.32, NS$. For parent education level and student ethnicity, respectively, the difference between groups was not significant, $x^2 (2, N = 9) \leq 1.56, NS$, as the values did not exceed the critical value for the 0.05 probability level (5.99).

**Setting**

Parents implemented the intervention in a one-to-one context (i.e., parent-child dyad) in their home. I instructed parents to select a non-distracting space in their home that was conducive to implementing the intervention (e.g., free of noise). The weekly direct
observation of the intervention session and the feedback session were conducted either in the home setting or in public settings (e.g., elementary school library), depending on parent preference.

**Measures**

**Treatment integrity.** I collected treatment integrity data via direct observation, audiotaped recordings, and permanent products. During each weekly observation session, I observed the parent and recorded implementation integrity of both the intervention’s primary protocol and the tips and reminders. I used audiotaped recordings and permanent products to capture treatment integrity data from the other two sessions conducted each week. When collecting data from audiotaped recordings I excluded tips and reminders from the overall calculation of the session integrity that could not be measured based upon listening to the tape and not directly observing the session. For example, one tip and reminder asked parents to cover the students’ reading passage before asking the retell question. In most cases I could not determine that this was done simply from listening to the audiotape.

**Core components.** I measured treatment integrity of the core components of the HELPS Program using the HELPS One-on-One Program: Implementation Protocol (see Appendix B). This document outlines each of the core components of the intervention. I recorded treatment integrity of the core components of the HELPS Program using the Observation Checklist for Implementing the HELPS One-on-One Program (see Appendix C). I calculated treatment integrity of the core components as the percentage of applicable steps completed correctly.
**Tips and reminders.** I recorded treatment integrity of the tips and reminders for implementing the HELPS Program using the Observation Checklist for Implementing the HELPS One-on-One Program. I calculated treatment integrity of the tips and reminders as the percentage of applicable steps completed correctly.

**Intervention session summary.** I used the Observation Summary Form (see Appendix D) to summarize the treatment integrity data recorded on the Observation Checklist for Implementing the HELPS One-on-One Program. On the Observation Summary Form I also recorded the length of the student intervention session, the length of the feedback session, a rating of the parent’s enthusiasm and a rating of the parent’s organization (on a scale of 1 to 5; 1=Poor, 3=Average, 5=Outstanding) while delivering the intervention, as well as any questions or concerns raised by the parent (see Appendix E for average session length, feedback length, rating of enthusiasm, and rating of organization for each case).

**Treatment acceptability.** I assessed parents’ acceptability of the HELPS Program using a modified version of the Intervention Rating Profile (IRP-15; Martens, Witt, Elliot, & Darveaux, 1985; see Appendix F). The IRP-15 is a 6-point Likert scale measure that assesses intervention acceptability. The reliability of the IRP-15 using Cronbach’s alpha is .98 (Martens et al.). I modified the language of the IRP-15 to be appropriate for administration to parents of children with reading problems, as was done by LeBlanc (2009).

I assessed parents’ judgments about the feedback procedures (i.e., performance feedback and self-evaluation) by administering a modified Intervention Rating Profile, as was done by DiGennaro et al. (2005). I duplicated the measure to create two forms, one for each condition, and I modified the language to reflect the two feedback conditions (see
Appendix G for a copy of the Performance Feedback form and Appendix H for a copy of the Self-Evaluation form). I also interviewed parents about their experience using the HELPS program and working with me to receive feedback. A copy of the interview questions is in Appendix I.

**Student performance.** I assessed student participants at their school or in their home pre- and post-intervention using the following measures: the Gray Oral Reading Test, Fifth Edition (GORT-5; Wiederholt & Bryant, 2001) and the Test of Word Reading Efficiency, Second Edition (TOWRE-2; Torgesen, Wagner, & Rashotte, 2012). I also administered the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999) during the pre-intervention assessment only. Prior to administering the pre-tests and post-tests I obtained verbal assent from each student participant to administer the assessment measures. I also collected students’ fall and winter Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Next Oral Reading Fluency (DORF, Good & Kaminski, 2011) benchmarking scores.

**Comprehensive Test of Phonological Processing (CTOPP).** The CTOPP is a normed measure of phonological processing skills. I administered two subtests from this measure only as pre-tests: (a) Rapid Letter Naming (RLN) and (b) Rapid Digit Naming (RDN). A composite score, Rapid Naming, was derived from performance on the RLN and RDN subtests. The RLN and RDN subtests were administered to screen students’ ability to recall phonological information from long-term memory and to determine if it was appropriate for each student participate in the study. The average coefficient alpha is .82 for the RLN subtest and .87 for the RDN subtest. The coefficient alpha of the Rapid Naming composite is .92 for
children ages 7 and older. The test-retest reliability of the RLN and RDN subtests for children ages 5 to 7 is .97 and .91 respectively, and for children ages 8-17 is .72 and .80 respectively (Wagner et al., 1999).

*Gray Oral Reading Test, Fifth Edition (GORT-5).* The GORT-5 is a normed measure of oral reading and comprehension. This measure is comprised of 16 developmentally sequenced reading passages, each followed by five comprehension questions. Administering the measure results in five scores: (a) Rate; (b) Accuracy; (c) Fluency; (d) Comprehension; and (e) Oral Reading Index. The Oral Reading Index is a composite score derived from combining the Fluency (i.e., Rate and Accuracy) and Comprehension subtests standard scores. The GORT-5 is considered a valid measure of reading ability, with a high degree of reliability. Average coefficient alphas for the four subtests and composite range from .92 to .96, and the average test-retest reliability coefficients range from .82 to .90 for the four subtests and is .87 for the Oral Reading Index. The average alternate forms reliability coefficients for the subtests range from .84 to .94 (Wiederholt & Bryant, 2012).

*Test of Word Reading Efficiency, Second Edition (TOWRE-2).* The TOWRE-2 is a nationally normed measure of word reading accuracy and fluency. This measure is composed of two subtests: (a) Sight Word Efficiency (SWE) and (b) Phonetic Decoding Efficiency (PDE). The SWE subtest measures the number of real printed words that can be accurately read in 45 seconds. The PDE subtest measures the number of pronounceable printed nonwords that can be accurately decoded in 45 seconds. The TOWRE-2 is considered a valid measure of word reading ability, and has average test-retest reliabilities that range from .87 to
.91 and average alternate forms reliability coefficients that range from .91 to .92 (Torgeson et al., 2012).

*Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Next- Oral Reading Fluency (DORF).* DORF is a measure of students’ advanced phonics and word attack skills, reading fluency and automaticity of connected text, and reading comprehension. This measure is designed for students in grades 1-6, and students are administered grade-level passages. This measure was used by the student participants’ school as a benchmarking measure, meaning that it was administered three times during the school year: fall, winter, and spring. For benchmarking purposes each student is administered three grade-level texts and asked to read aloud for one-minute. While the student reads aloud, the test administrator records any reading errors, such as omissions, substitutions, and hesitations of 3 seconds or more. Self-corrected words are marked as correct. A student’s DORF score is the median Words Correct Per Minute (WCPM) score of the three passages. There is also a Passage Retell component to the DORF measure; however, the Retell scoring is separate from the DORF WCPM score.

The two-week alternate form reliability for three-passage groups of DORF passages, based on students’ median WCPM score, is .97 for second grade passages and .96 for third and fourth grade passages. The two-week test-retest reliability for three-passage groups of DORF passages, based on students’ median WCPM score, is .91 for second grade, .93 for third grade, and .97 for fourth grade. Inter-rater reliability of the three-passage groups of DORF passages is .99 for grades 2 – 4 (Good, Kaminski, Dewey, Wallin, Powell-Smith, & Latimer, 2013).
Criterion-related validity was calculated based on the comparison of the DORF measure with the Group Reading Assessment and Diagnostic Evaluation (GRADE), a norm-referenced test of reading achievement. Predictive validity coefficients are .69 (beginning of year) and .76 (middle of year) for second grade, .66 (beginning of year) and .67 (middle of year) for third grade and .77 (beginning and middle of year) for fourth grade. Concurrent validity coefficients, based on comparison with the Standard 4th Grade Reading Passage used in the NAEP 2002 Special Study of Oral Reading, is .91 for second grade, .96 for third grade, and .89 for fourth grade (Good et al., 2013).

**Procedures**

**Student participant recruitment.** I used four criteria to determine eligibility for student participation in this study. First, students needed to have a DORF score on the fall benchmark passages of the DIBELS Next (Good & Kaminski, 2011) that fell within the 10th to 50th percentile. Second, the reading teachers nominated students with low benchmark scores for participation in this study as at-risk or struggling students. That is, the reading teachers determined that the student would likely benefit from a reading intervention targeting reading fluency. Third, students could not be receiving special education services or English as a Second Language services. Fourth, parents of students eligible for participation provided informed consent and agreed to the three conditions outlined for parent participants (see next section). The first ten students I identified who met these criteria were included in this study; however, more than ten parents were interested in participating in the study. Thus, a total of 12 student participants were initially included in the study.
**Parent participant recruitment.** I invited 112 parents of students who met the first three selection criteria to participate in this study via a letter sent to them during the 2013/2014 academic school year. In this letter I invited parents to an informational session about the reading intervention and the study. Of the 112 parents, 33 (29%) returned the completed form indicating interest in learning more about the study, and one parent returned the slip on which she stated that she had already been trained in the reading intervention. A total of 17 parents attended one of two informational meetings held to provide additional information about the study, answer study-related questions, and distribute the informed consent form and background questionnaire. I also attended the school’s Open House, held in the evening, where I provided additional information about the study, answered study-related questions, and distributed the informed consent form and background questionnaire to parents who had received a letter.

Parents who were interested in the intervention and committed to participating in the study provided written consent for themselves and their child to participate in the study. Parents agreed to the three conditions outlined for participation in the study: (a) attend one training workshop to learn about the intervention; (b) implement the intervention three times per week for 8-10 weeks; and (c) meet with me weekly to allow me to observe a session, where feedback would be provided in different forms or not at all. A total of 12 parents submitted a completed informed consent form to participate in the study. Prior to the parent training one parent was unable to continue with the study due to the family moving out of the county. Eleven parent/child dyads began the study and participated in the baseline phase and first experimental feedback phase; however, two dyads did not continue the study long
enough to participate in the second experimental feedback phase. Therefore, a total of 9 parent/child dyads fully participated in this study.

**Parent training.** Prior to implementing the reading program with their child, all parents attended a 3-hour training workshop. Six parents attended one training workshop together and three parents attended another training workshop together. The two workshops were held within a week of each other. The purpose of the workshop was to familiarize parents with evidence-based reading fluency procedures and train the parents on the reading intervention they used with their child. During this workshop I provided parents with information about reading fluency, the purpose of the research project, previous research that was conducted on the intervention, and how to implement each component of the intervention. I reviewed each component of the intervention in depth, providing a demonstration of each step using the intervention materials. When sharing information about the purpose of the research project, I told parents that they would receive two forms of feedback during the course of this study but did not share any information about the perceived effectiveness of either form of feedback.

During the training workshop I provided parents with all of the intervention materials, including the intervention training manual, teacher reading passages, student reading passages, and student folder, where data would be recorded on the student graph, star chart, and Progress Tracking Form. I also provided a reward box with prizes, and an audiotape player and blank tapes for recording all intervention sessions, including instructions for using the audiotape player. At the completion of the workshop, I gave parents an intervention implementation quiz to complete before leaving. The quiz consisted of 12 True/False, 11
multiple choice, and three short answer/applied questions (see Appendix J). Prior to dismissing parents from the workshop, I reviewed and discussed the correct answers to the questions to ensure that parents understood the correct implementation procedures.

**Overview of HELPS Program reading intervention.** The Helping Early Literacy with Practice Strategies (HELPS) Program is designed to strengthen students’ reading fluency (Begeny, 2009). HELPS is a research-based intervention that incorporates several evidence-based teaching strategies, including repeated reading, modeling, phrase-drill error correction, goal-setting, performance feedback, and a motivational/reward procedure. HELPS was designed to be used with students in all primary grades. The HELPS Program has a strong research base supporting its use for improving reading fluency and comprehension (e.g., Begeny, Mitchell, Whitehouse, Harris, & Stage, 2011; Begeny, Laugle, Krouse, Lynn, Tayrose, & Stage, 2010; Mitchell & Begeny, 2014).

Two primary protocols provide guidance for implementing the HELPS Program: (a) the Scripted Directions protocol and (b) the Tips and Reminders for Implementation. The Scripted Directions protocol (see Appendix K) provides the parent with a verbal script for administering each component of the HELPS Program to their child. The script is interspersed with directions the parent should follow for implementing each component of the HELPS Program. The Tips and Reminders for Implementing the Program protocol (see Appendix L) provides 31 recommendations and procedural reminders the parent should follow when administering the intervention. These tips and reminders not only include tips and reminders about the primary procedures of the HELPS Program, such as covering the reading passage while the student completes the retell check, but also include tips and
reminders about teaching behaviors, such as providing genuine praise to the student. Implementing the tips and reminders is not considered optional (Begeny, 2009).

The following materials are needed to implement each HELPS session: (a) Scripted Implementation Protocol form; (b) Tips and Reminders for Implementing the Program protocol form; (c) stop watch; (d) parent passages; (e) student passages; (f) dry-erase marker; (g) pencil; (h) student’s progress tracking form (see Appendix M for an example); (i) student’s graph (see Appendix N); (j) student’s star chart (see Appendix O); (k) bonus bag; (l) implementation flow chart (see Appendix P); and (m) prize box. The implementation flow chart is a quick-reference guide for sequencing the core procedures of the HELPS Program during each session.

To implement a HELPS session, the parent first reads an introductory statement that includes reading expectations (e.g., read quickly & with good expression) to the child. The parent then times her child with a stopwatch while her child reads Passage A. Immediately following this first oral reading the parent instructs her child to re-tell the story. If the child has met his reading goal based on Passage A performance and the re-tell check, the parent delivers praise to her child and graphs her child’s performance. The parent then administers a timed reading using Passage B. Immediately following this timed reading, the parent identifies words that were pronounced incorrectly and/or read less fluently, and instructs her child to practice reading these words within selected phrases three times. This process is called the phrase/drill procedure. The parent then times her child re-reading Passage B. Following the child’s second reading of Passage B, the parent then models fluent oral reading of the passage to her child. The parent then times her child re-reading Passage B. After this
final reading, the parent graphs the first and third timed reading of Passage B and provides praise and feedback to her child. The parent then typically awards two stars on her child’s star chart, one for initially meeting the reading goal and one for putting forth good effort during the rest of the session. In the next step, the parent records her child’s data on the Progress Tracking Form. Lastly, the parent reviews the implementation protocol and records any missing/incomplete steps on the Progress Tracking Form.

If, after completing the reading of Passage A and the re-tell check, the child did not meet his reading goal, the structure/sequence of the remaining session is slightly different from the condition in which the child met his goal. After the first reading of Passage A, the parent models fluent oral reading of Passage A. The parent then times her child while he re-reads Passage A. Following this step, the parent uses the phrase/drill procedure, as described in the preceding paragraph. The parent then times her child re-reading Passage A for the third time. The parent then uses the phrase/drill procedure again. Following this last phrase-drill procedure, the parent graphs her child’s first and third timed reading performance on Passage A. The parent then awards one star on the star chart for working hard, records student data on the Progress Tracking Form, and records any missing/incomplete intervention steps on the Progress Tracking Form.

A HELPS session is composed of 13 steps if, after the timed reading of Passage A and the retell check, the child meets his reading goal, and 12 steps if the child does not meet his reading goal. The last step of this intervention (reviewing the flowchart to determine any missing/incomplete steps), however, was not included in the training and program materials for the parent participants. Given that this study aimed to evaluate the effects of feedback on
implementation integrity, I did not want to include an intervention step that potentially provided parents with feedback that was in addition to the feedback provided in either of the two feedback phases, or provided parents with feedback when in the Baseline phase. Therefore, if the child met his reading goal there were a total of 12 steps and if the child did not meet his reading goal there was a total of 11 steps.

**Intervention.** I instructed parents to implement the intervention three times per week (e.g., every Monday, Wednesday, and Friday) for approximately 10-12 minutes per session. The total number of sessions each parent completed ranged from 14-31 ($M = 23.78$, $SD = 5.36$) across 8-10 weeks. I required parents to audiotape each intervention session, and I provided oral and written instructions for using the audiotape during the training workshops. I assessed intervention implementation levels during each phase of the study. Four parents had at least one session that was not recorded, with one parent not recording four sessions, one parent not recording two sessions, and two parents not recording one session.

**Phases.** Participants proceeded through three phases: (a) baseline, (b) self-evaluative feedback, and (c) performance feedback. The latter two phases were counter-balanced across participants, such that five parents completed self-evaluation as the first feedback method and I provided four parents performance feedback as the first feedback method.

**Implementation baseline.** Following the training, parents began implementing the intervention with their child in their home. I conducted the first observation session within one week of the parent-training workshop, and audiotaped each observation session. While directly observing the parent, I followed the Steps and Guidelines for Observing Parents During Their Implementation of the HELPS One-on-One Program: Baseline Phase protocol
(see Appendix Q), following eleven steps. These steps included (a) recording the identifying information on the Observation Summary Form (OSF), (b) waiting patiently until the parent was ready to begin the session, (c) starting the stopwatch to time the duration of the session, (d) recording the intervention steps completed correctly, (e) following along with the examiner copy of the reading passage and recording student reading errors, (f) recording the session data on the OSF, (g) recording Inter-Scorer Reliability Agreement data, (h) resetting the stopwatch, asking the parent if he/she had any questions about the session, and recording the questions but not providing any feedback, (i) thanking the parent for his/her time and effort, (j) stopping the stopwatch and recording applicable data to the OSF, and (k) reviewing all applicable steps on the observation protocol to ensure that I completed the observation correctly.

I recorded observation data on the Observation Checklist for Implementing the HELPS Program and the Observation Summary Form. During the baseline phase parents did not receive assistance or feedback from me. Participants moved into the next phase after a minimum of three consecutive intervention sessions and at least one observation session.

**Self-evaluative feedback.** During each direct observation of parent participants, I audiotaped the parent delivering the intervention to the student participant. I followed the Steps and Guidelines for Observing Parents During Their Implementation of the HELPS One-on-One Program: Self-Evaluation Phase protocol (see Appendix R) for observing implementation of the intervention, following thirteen steps. While observing the parent, I completed the Observation Checklist for Implementing the HELPS Program.
After the parent executed the intervention protocol I immediately rewound and played the audiotaped session for the parent while the parent completed the [Parent’s] Observation Checklist for Implementing the HELPS Program (see Appendix S). This form is identical to the Observation Checklist I used in all observations with the exception that the language was changed such that “I” replaced “parent” and “my child” replaced “student.” I then provided time for the parent to review the parent-completed Observation Checklist, and then collected the Observation Checklist from the parent. I did not provide any feedback to the parent. If parents had questions or concerns, I recorded this information on the Observation Summary Form but did not provide direct responses to the questions and/or concerns; rather, I told parents that their questions and/or concerns were recorded and would be discussed when they were able to receive performance feedback. I also recorded the amount of time needed for the parent to listen to the audiotaped session, complete the Observation Checklist, and review the Checklist. If this was the first feedback phase, participants moved into the performance feedback phase after a minimum of three sessions and at least one observation session.

I compared my ratings of the parents’ implementation of the core components and the tips and reminders to the parents’ self-ratings of their implementation of the core components and the tips and reminders by calculating percent agreement. I calculated percent agreement by dividing the number of agreements by agreements plus disagreements, multiplied by 100. Percent agreement was calculated separately for core components and for the tips and reminders. Across sessions and participants, the average agreement for the core components was 90% (range = 55% - 100%; SD = 12), and the average agreement for the tips and reminders was 75% (range = 48% - 94%; SD = 11). In almost all cases, the disagreement in
ratings was a result of parents endorsing that they completed steps of the program that I did not observe in the session. On average, parents endorsed that they implemented 92% of the core components and 89% of the tips and reminders, while I endorsed that parents implemented 85% of the core components and 73% of the tips and reminders.

**Performance feedback.** During each direct observation of parent participants, I audiotaped the parent delivering the intervention to the student participant. While observing I adhered to the Steps and Guidelines for Observing Parents During Their Implementation of the HELPS One-on-One Program: Performance Feedback Phase protocol (see Appendix T), following twenty steps. In addition to the same observation steps that were followed in the Baseline and Self-Evaluation phases, the performance feedback phase observations also included (a) immediate corrective feedback to the parent when a core step of the intervention protocol was missed; (b) praising and supporting the parent after the parent completed the intervention session; (c) reviewing session data (i.e., duration of session, number of core steps completed correctly and incorrectly, number of tips and reminders completed correctly and incorrectly, evaluation of parent’s enthusiasm & organization) (d) reviewing implementation errors of the core components and answering any questions posed by the parent, (e) reviewing implementation errors of the tips and reminders and answering any questions posed by the parent, (f) reviewing and explaining parent’s enthusiasm and organization ratings (if not rated Outstanding), (g) providing additional feedback that may have assisted the parent in implementing the intervention (e.g., tips for organizing materials), (h) asking the parent if he/she had additional questions about HELPS implementation procedures, and (i) identifying, and possibly demonstrating, 1-3 things the parent should
improve upon for upcoming sessions. I recorded data from each direct observation on the Observation Checklist for Implementing the HELPS Program and the Observation Summary Form. If this was the first feedback phase, participants moved into the self-evaluation phase when treatment integrity decreased or was stable; however, a minimum of three consecutive intervention sessions was required before moving into the next phase.

During the baseline phase I observed parents an average of 1.56 times (range = 1-3, \(SD = 0.88\)). During the self-evaluation phase I observed parents and asked them to conduct the self-evaluation feedback method an average of 2.67 times (range = 2-3, \(SD = 0.50\)). During the performance feedback phase I observed parents and provided performance feedback an average of 3.67 times (range = 2-5, \(SD = 0.87\)).

**Inter-rater reliability and procedural integrity**

A total of three research assistants (RAs) assisted with the data coding for this study. I recruited undergraduate research assistants who were already trained in the HELPS Program and had experience implementing the intervention with students as well as documenting implementation integrity. Two RAs assisted with coding 20\% of recorded parent intervention sessions to calculate inter-rater reliability. I compared my ratings of the parents’ implementation of the core components and the tips and reminders to an RA’s ratings of parents’ implementation of the core components and the tips and reminders by calculating percent agreement. I calculated percent agreement by dividing the number of agreements by agreements plus disagreements, multiplied by 100. Percent agreement was calculated separately for core components and for the tips and reminders. Across sessions and
participants, the average agreement for the core components was 95% (range = 83% - 100%),
and the average agreement for the tips and reminders was 76% (range = 45% - 90%).

I collected procedural integrity data on my procedures for observing the intervention
sessions and following the protocol specific to each phase (i.e., the degree to which I did
what I said I would do and do so accurately). One RA coded 94% of the observation sessions
I conducted with the parents. The second coding by the RA was done to calculate inter-rater
reliability of the procedural integrity, that is, the agreement between the ratings I gave myself
for accurately adhering to the observation protocol that was specific to each of the phases and
the ratings given by the RA.

My overall self-rated procedural integrity for correctly observing, collecting
participant implementation data, and adhering to the procedures specific to each phase of the
study (e.g., providing vocal feedback only during the performance feedback phase) was 99%.
The step that I occasionally omitted during the feedback sessions was thanking the parent for
his/her time. A research assistant also evaluated my procedural integrity for 66 of the 70
(94%) observation/feedback sessions by reviewing audiotapes of these sessions. I calculated
inter-rater reliability as the number of instances of agreement divided by agreements plus
disagreements, multiplied by 100. The inter-rater reliability between my self-ratings and the
ratings of the research assistant for the procedural integrity was 97%. Overall, I demonstrated
high adherence to the observation protocols that were specific to each phase of the study.

Prerequisite Analyses

Before I moved to compare data in the three phases for the purpose of answering my
research questions and hypotheses, I analyzed the data depicting parents’ treatment integrity
using visual inspection, including completing a number of computations to ensure that subsequent comparisons via data analysis were valid. Treatment integrity data from each intervention session was plotted on two line graphs: (a) integrity of the core intervention components and (b) integrity of the tips and reminders. On each graph the vertical axis represents the integrity level, reported in the form of a percentage, and is scaled from 0 to 100 with equally spaced tic marks. The horizontal axis represents the session number.

For each participant I examined the implementation integrity data of the core components and the tips and reminders separately. I analyzed the data within each of the conditions to examine level and trend, and to determine if level and trend were stable, and then compared data in two selected conditions on five possible variables: (1) the change in level of implementation integrity from one phase to the compared phase, (2) the change in trend from one phase to the compared phase, (3) if there was an immediate change in level of implementation integrity after a phase change (only compared adjacent conditions), (4) if there was a delayed change in implementation integrity following a phase change (only compared adjacent conditions), and (5) the percent of nonoverlapping data. The following sections outline the criteria I applied to describe data within and between compared/adjacent phases.

**Within-condition analyses.** All cases had three phases: baseline, self-evaluation, and performance feedback. For each phase, I examined the overall level, level stability, trend direction, and trend stability.

**Overall level.** I assessed the overall level of implementation integrity of the core components and the overall level of the tips and reminders within each phase by calculating
median implementation levels. Calculating the median value, as opposed to the mean value, is recommended because the mean is influenced by extreme values (Gast & Spriggs, 2010). To determine the median integrity level I first sequenced the data point values from lowest to highest. For an odd number of data points, the median is the middle value; for an even number of data points the median is the average of the two middle values.

**Level stability.** To assess the stability of the data within each phase, I first used the median value previously calculated to draw a median line parallel to the abscissa. I then drew a stability envelope around the median line. The stability envelope consisted of a 20% band that was placed parallel to the median line, and shifted up or down to capture the greatest number of data points. The distance between the two parallel lines indicated how much variability there could be in the data set to be considered stable. I used the 80%-20% criteria, whereby if 80% of the data points fell on or within the 20% band around the median line, the data were considered stable, a criterion commonly used among applied researchers (Gast & Spriggs, 2010). If the level was stable, I concluded that the data may be compared with the level from another phase (that was also stable). If the level was not stable, I recorded “IC” for Invalid Comparison, and did not make any subsequent comparisons within the case using that feature.

**Trend direction.** To assess the slope of the data in the baseline and two experimental phases, I constructed trend lines using the split-middle method (Gast & Spriggs, 2010; Tawney & Gast, 1984). The split-middle method is composed of four steps. First, I divided the data to be summarized into two equal parts. Second, I found the intersections of the mid-rate and mid-date for each half. The mid-rate is the middle data point when counting up or
down the data path; if there is an even number of data points, the mid-rate is between the two middle data points. The mid-date is the middle data point when counting left on the data path; if there is an even number of data points, the mid-date is between the two middle data points. Third, I drew a line through the data which passed through both of the intersections found in the previous step. Finally, if the number of data points on or above the line was not the same as the number of data points on or below the line, I moved the line up or down (keeping it parallel to the original line) until a balance was achieved. The split-middle method provided a better portrayal of the overall trend in data (Richards, Taylor, Ramasamy, & Richards, 1999), as this method is considered a more reliable estimate of trend (Tawney & Gast). I assessed the trend, or slope, of the data within each condition to determine if it was accelerating (increasing in ordinate value over time), decelerating (decreasing in ordinate value over time), or zero accelerating (data series is parallel to the abscissa). I also recorded whether the direction of the trend was improving or deteriorating.

**Trend stability.** I evaluated trend stability by using the same 80%-20% criteria (Gast & Spriggs, 2010) that I used when I evaluated level stability. I placed a stability envelope around the trend line, making sure that both sides of the stability envelope remained parallel to the trend line. I determined that the trend was stable if 80% of the data points fell on or within the stability envelope. If the trend line was stable I concluded that the data may be compared with a stable trend line of another phase. If the trend was not stable, I recorded “IC” for Invalid Comparison, and did not make any subsequent comparisons within the case using that feature.

**Between-condition analyses.** For each participant I compared the level and trend of
the three conditions, if the levels and trends were sufficiently stable to compare, as discussed in the previous section. For each participant I also analyzed the data in adjacent conditions to determine the absolute level change and relative level change. Lastly, I calculated the percentage of non-overlapping data points.

**Absolute level change.** Calculating the absolute level change between conditions provides an indication of the immediate strength of the intervention on the dependent variable (Gast & Spriggs, 2010). I used the following steps for calculating an absolute level change between adjacent conditions. First, I identified the last data point in the first phase and the first data point in the subsequent phase. Second, I subtracted the data point with the smaller value from the data point with the larger value. Third, I noted whether the change in level improved or deteriorated.

**Relative level change.** The calculated relative level change provided an indication of whether there was a change in the dependent variable after introducing the intervention. Unlike the absolute level change, which indicated whether there was an immediate change in the dependent variable after introducing the intervention, the relative level change may detect delayed effects (Gast & Spriggs, 2010). To calculate the relative level change I first identified the median value in the last half of the data series in the first condition. Second, I identified the median value in the first half of the data series in the second condition. Third, I subtracted the smaller value data point from the larger value data point. Finally, I recorded whether the level change indicated improvement or deterioration.

**Change in trend direction.** I compared the trend lines of the conditions. I recorded any changes in the trend line from the first phase to the second phase into one of the
following categories: (a) zero celerating to accelerating, (b) zero celerating to decelerating, (c) accelerating to decelerating, (d) decelerating to accelerating, and (e) zero celerating to zero celerating (at higher or lower level). For cases where the first experimental feedback phase trend line was compared to the second experimental feedback phase trend line, I determined that treatment integrity was higher in the second feedback phase than the first feedback phase if I observed any of the following conditions: (a) change from zero celerating trend in first feedback phase to accelerating trend in second feedback phase, (b) change in trend from decelerating trend in first feedback phase to accelerating trend in second feedback phase, (c) change in trend from decelerating trend in first feedback phase to an immediate increase in level with zero celeration at the increased level in the second feedback phase, or (d) change in trend from zero celeration to an immediate increased level change and zero celeration at the increased level.

For cases where the baseline trend line was compared to either the self-evaluation phase trend line or performance feedback trend line, I determined that treatment integrity was higher in the self-evaluation or performance feedback phase than the baseline phase if I observed any of the following conditions: (a) change from zero celerating trend in baseline phase (at a decreased level) to accelerating trend in self-evaluation or performance feedback phase, (b) change in trend from decelerating trend in baseline phase to accelerating trend in self-evaluation or performance feedback phase, (c) change in trend from decelerating trend in baseline phase to an immediate increase in level with zero celeration at the increased level in the self-evaluation or performance feedback phase or (d) change in trend from zero celerating trend in baseline phase to zero celerating trend at an increased level in self-evaluation or
performance feedback phase.

**Percentage of non-overlapping data.** I calculated the percentage of non-overlapping data point values (PND) to determine the impact of the intervention on treatment integrity. To calculate the PND I (a) determined the range of data point values in the first condition; (b) counted the number of data points plotted in the second condition; (c) counted the number of data points from the second condition that fell outside the range of values of data points in the first condition and; (d) divided the number of data points that fell above the range of the first condition by the total number of data points in the second condition and multiplied this number by 100 (Scruggs, Mastropieri, & Casto, 1987). I used a cut point of 70%, based on Scrugg’s and Mastropieri’s (1998) guidelines, whereby I determined that there was a difference between phases if there was 70% or greater non-overlapping data points.

**Data analysis procedures.** To answer my two research questions and four hypotheses, I first conducted a visual analysis of each case, separately analyzing the treatment integrity of the core components and the tips and reminders. For each case I entered all of the features of the visual analysis into a table. Given that there are multiple ways of examining single-case data using visual analysis, and there are no set criteria for statistical analysis (Kratochwill, et al., 2012), I incorporated the steps and features of analyzing the outcome variable (i.e., treatment integrity) outlined by Kratochwill and colleagues as part of the *Standards* for single-case design. I first examined the baseline phase of each case to determine whether the data were stable, first with respect to level, then with respect to trend. Second, I examined data from each of the experimental feedback phases to determine whether the level was stable, and whether the trend was stable. Third, I compared
data in two selected phases. For each comparison (i.e., baseline to self-evaluation, baseline to performance feedback, and self-evaluation to performance feedback) I examined the two phases on several features. Each comparison was recorded in the table as a directional statement (e.g., SE > Baseline, Performance > SE). If a phase did not have either a level or trend that was sufficiently stable to compare with another phase, I considered the data for that particular phase to be invalid and excluded the data from any further within-case comparisons with other phases.

The features on which two phases were compared included: (a) level, (b) trend, (c) data stability (i.e., level stability and trend stability), (d) absolute level change, (e) relative level change, and (f) percent non-overlapping data. If the levels of two compared conditions were sufficiently stable to compare, I recorded which of the two conditions demonstrated an integrity level that was higher than the other. If there was no difference in level, I recorded this as equal, and I treated these data as a tie. Next, if the trend lines of two compared conditions were sufficiently stable to be compared, I recorded which of the two phases demonstrated a trend line that demonstrated improved integrity when compared to the other (see description of trend comparisons on pages 48-49). If there was no difference in trend, I recorded this as equal, and I treated these data as a tie. If the level or trend was insufficiently stable to compare, I recorded this within the table as an invalid comparison. I calculated the absolute level change and relative level change only for adjacent conditions, and recorded the phase that demonstrated an improved absolute and the phase that demonstrated an improved relative level change. For nonadjacent conditions I recorded “NC” (Not Comparable) in the cells of the table. Lastly, I calculated the percent of non-overlapping data (PND). Although
calculating PND is typically only done on data in adjacent phases, I calculated it on the baseline and phase 3 data because there was only one baseline phase in the study in which to compare the two experimental feedback phases, and the comparison provided a means of exploring the degree of overlap between these two phases.

At the bottom of each table I recorded the number of features that supported, did not support, or were tied (i.e., no clear difference between the compared phases), written as inequalities (i.e., Self-Evaluation (SE) > Baseline, Performance Feedback (PF) > Baseline, PF > SE). I then determined which of the two compared conditions in each case demonstrated greater treatment integrity by looking at whether a majority of valid comparisons supported or disputed the inequality. If a case had the majority of the valid comparisons fall within the “Tie” column, or if the valid comparisons were dispersed among all three of the categories, with neither the “support” or “did not support” receiving a majority, I concluded neither phase was superior to the other, and labeled this decision “tie.” If a case had invalid data (both level and trend were insufficiently stable to compare), I recorded this decision in the table as “Invalid Comparison” (IC) and I excluded the data from the overall across case comparison of the two phases.
Results

Visual Analysis of the HELPS Core Components Treatment Integrity

The percentage of core component steps correctly implemented each session by parent participants is presented in Figures 1 and 2. Data points that are not filled represent sessions in which I directly observed the parent implementing the HELPS Program, and depending on the phase, where I instructed the parent to complete the self-evaluation procedures or I provided performance feedback.

Parents in cases A through E, where self-evaluation was the first feedback method, demonstrated varying integrity levels with the introduction of self-evaluation. The median integrity level and range of each phase for each case is reported in Table 1.

Table 1  
Median Integrity Level and Range of Core Components

<table>
<thead>
<tr>
<th>ABC Cohort</th>
<th>Baseline</th>
<th>Self-Evaluation</th>
<th>Performance Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case A</td>
<td>68.5%</td>
<td>91%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>(55-82)</td>
<td>(82-100)</td>
<td>(91-100)</td>
</tr>
<tr>
<td>Case B</td>
<td>18%</td>
<td>36%</td>
<td>36%</td>
</tr>
<tr>
<td></td>
<td>(9-27)</td>
<td>(27-36)</td>
<td>(27-45)</td>
</tr>
<tr>
<td>Case C</td>
<td>82.5%</td>
<td>91%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>(73-100)</td>
<td>(64-92)</td>
<td>(73-100)</td>
</tr>
<tr>
<td>Case D</td>
<td>92%</td>
<td>83%</td>
<td>91%</td>
</tr>
<tr>
<td></td>
<td>(73-100)</td>
<td>(25-100)</td>
<td>(82-100)</td>
</tr>
<tr>
<td>Case E</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>(91-100)</td>
<td>(100-100)</td>
<td>(91-100)</td>
</tr>
</tbody>
</table>
In two cases (Cases A & C) following an accelerating baseline trend line, the introduction of self-evaluation resulted in an absolute level increase for Case A and an absolute level decrease for Case C. In both cases, however, the relative level change increased with the introduction of self-evaluation. Case B had a stable decelerating baseline trend with an immediate improved level change (from 18% to 36%) following the introduction of self-evaluation. The data trend, however, indicated a decline over time despite having an improved level when compared to baseline. In Case D the parent had a stable zero-celerating level in the baseline phase that decreased and became unstable with the introduction of self-evaluation. In one case (Case E) the parent began implementing the program with 100% integrity and maintained this level across both feedback phases.

The introduction of performance feedback to Cases A through E, suggests increased
Figure 1. Treatment integrity of HELPS Program core components for participants in ABC cohort. ♦ = session not directly observed; ♦ = direct observation/feedback for Phases 2 & 3
Figure 2. Treatment integrity of HELPS Program core components for participants in ACB cohort. ♦ = session not directly observed; ◇ = direct observation/feedback for Phases 2 & 3
treatment integrity levels for three cases (Cases A, C, & D). The introduction of performance feedback to Case B showed an immediate positive level change; however the overall level remained the same as that in the self-evaluation phase and the data trend declined, as it was in both the baseline and self-evaluation phases. For Case E the level remained the same as in the self-evaluation phase, at 100%.

The second cohort, Cases F through I, received performance feedback first followed by self-evaluation. Three of the four cases (Cases F, H, & I) demonstrated improving treatment integrity in the baseline phase, as demonstrated with stable accelerating trend lines. Following the introduction of performance feedback with these three cases, treatment integrity levels continued to improve, although the level and trend lines for Case H were unstable. Cases F and H improved the level to 100% integrity during this phase. Case G was the only case that demonstrated deteriorating treatment integrity during the baseline phase. The introduction of performance feedback showed an immediate level increase (from 73% to 91%) and an accelerating trend line.

With the introduction of self-evaluation to Cases F through I, the treatment integrity generally maintained high integrity or improved integrity compared to the performance feedback phase. In Case G the median improved to 100% for the entire phase. In Case H the median improved and the data reflected a continued accelerating trend line. In Case I the median remained at 100% but there was no fluctuation in the data, as there was in the performance feedback phase. Case F was the only case to differ. Although the median integrity level remained at 100%, the data reflected a decelerating trend line.
Summary of the visual analysis of core components. Examining data across all cases, two of the nine cases had stable decelerating baseline trend lines (Cases B & G). The introduction of a feedback method (Self-evaluation for Case B and Performance Feedback for Case G) was reflected in the data as an immediate improved level change. For both cases, treatment integrity level in phases two and three remained higher than the baseline level. In five of the nine cases (Cases A, C, F, H, & I) participants had stable accelerating baseline phases, demonstrating improved treatment integrity with the absence of structured feedback. In all cases, the introduction of a feedback method had mixed results. Neither feedback method produced an immediate level change across all cases. Regardless of the order of the feedback sequence (i.e., either self-evaluation as the first or second feedback phase), across cases there was a general trend of treatment integrity improvement over time, as measured by the median integrity level of each phase.

The results of the visual analysis of the core components for each case are summarized in Tables 2 through 10.
Table 2

*Case A – Visual Analysis of Core Components*

<table>
<thead>
<tr>
<th></th>
<th>Stable/Trend Dir</th>
<th>Baseline vs Self-Eval</th>
<th>Baseline vs PF</th>
<th>PF vs SE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong></td>
<td>Level</td>
<td>N (68.5%)</td>
<td>IC</td>
<td>IC</td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>Y (accelerating)</td>
<td>B = SE</td>
<td>B = PF</td>
</tr>
<tr>
<td><strong>Self-Evaluation (SE)</strong></td>
<td>Level</td>
<td>Y (91%)</td>
<td>IC</td>
<td>PF &gt; SE</td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>Y (zero accelerating)</td>
<td>B = SE</td>
<td>PF &gt; SE</td>
</tr>
<tr>
<td><strong>Performance Feedback (PF)</strong></td>
<td>Level</td>
<td>Y (100%)</td>
<td>IC</td>
<td>PF &gt; SE</td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>Y (accelerating)</td>
<td>B = PF</td>
<td>PF &gt; SE</td>
</tr>
<tr>
<td><strong>Absolute Level Change</strong></td>
<td></td>
<td></td>
<td>B &lt; SE</td>
<td>NC</td>
</tr>
<tr>
<td><strong>Relative Level Change</strong></td>
<td></td>
<td></td>
<td>B &lt; SE</td>
<td>NC</td>
</tr>
<tr>
<td><strong>% Non Overlapping Data</strong></td>
<td></td>
<td></td>
<td>B &lt; SE</td>
<td>B &lt; PF</td>
</tr>
</tbody>
</table>

| Tallied Features |
|------------------|------------------|------------------|------------------|------------------|
| **Comparison**   | **N Valid Features** | **Supported** | **Not Supported** | **Tied** | **Decision** |
| SE > Baseline    | 4                | 3                | 0                | 0                | Y               |
| PF > Baseline    | 2                | 1                | 0                | 1                | T               |
| PF > SE          | 5                | 3                | 0                | 1                | Y               |

*Note: B = Baseline; PF = Performance Feedback; SE = Self-Evaluation; IC = Invalid Comparison; NC = Not Comparable Features; Y = Yes; N = No; T = Tied*
Table 3

*Case B – Visual Analysis of Core Components*

<table>
<thead>
<tr>
<th></th>
<th>Stable/ Trend Dir</th>
<th>Baseline vs Self-Eval</th>
<th>Baseline vs PF</th>
<th>PF vs SE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong></td>
<td>Level</td>
<td>Y (18%)</td>
<td>B &lt; SE</td>
<td>B &lt; PF</td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>Y (decelerating)</td>
<td>B = SE</td>
<td>B = PF</td>
</tr>
<tr>
<td><strong>Self-Evaluation (SE)</strong></td>
<td>Level</td>
<td>Y (36%)</td>
<td>B &lt; SE</td>
<td>PF = SE</td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>Y (decelerating)</td>
<td>B = SE</td>
<td>PF = SE</td>
</tr>
<tr>
<td><strong>Performance Feedback (PF)</strong></td>
<td>Level</td>
<td>Y (36%)</td>
<td>B &lt; PF</td>
<td>PF = SE</td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>Y (decelerating)</td>
<td>B = PF</td>
<td>PF = SE</td>
</tr>
<tr>
<td><strong>Absolute Level Change</strong></td>
<td></td>
<td></td>
<td>B &lt; SE</td>
<td>NC</td>
</tr>
<tr>
<td><strong>Relative Level Change</strong></td>
<td></td>
<td></td>
<td>B &lt; SE</td>
<td>NC</td>
</tr>
<tr>
<td><strong>% Non Overlapping Data</strong></td>
<td></td>
<td></td>
<td>B &lt; SE</td>
<td>B = PF</td>
</tr>
</tbody>
</table>

**Tallied Features**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>N Valid Features</th>
<th>Supported</th>
<th>Not Supported</th>
<th>Tied</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE &gt; Baseline</td>
<td>5</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>Y</td>
</tr>
<tr>
<td>PF &gt; Baseline</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>T</td>
</tr>
<tr>
<td>PF &gt; SE</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>T</td>
</tr>
</tbody>
</table>

*Note.* B = Baseline; PF = Performance Feedback; SE = Self-Evaluation; NC = Not Comparable Features; Y = Yes; N = No; T = Tied
Table 4

*Case C – Visual Analysis of Core Components*

<table>
<thead>
<tr>
<th></th>
<th>Stable/ Trend Dir</th>
<th>Baseline vs Self-Eval</th>
<th>Baseline vs PF</th>
<th>PF vs SE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong></td>
<td>Level</td>
<td>Y (82.5%)</td>
<td>B &lt; SE</td>
<td>B &lt; PF</td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>Y (accelerating)</td>
<td>B = SE</td>
<td>B = PF</td>
</tr>
<tr>
<td><strong>Self-Evaluation (SE)</strong></td>
<td>Level</td>
<td>Y (91%)</td>
<td>B &lt; SE</td>
<td>PF &gt; SE</td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>Y (zero accelerating)</td>
<td>B = SE</td>
<td>PF &gt; SE</td>
</tr>
<tr>
<td><strong>Performance Feedback (PF)</strong></td>
<td>Level</td>
<td>Y (100%)</td>
<td>B &lt; PF</td>
<td>PF &gt; SE</td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>Y (zero accelerating)</td>
<td>B = PF</td>
<td>PF &gt; SE</td>
</tr>
<tr>
<td><strong>Absolute Level Change</strong></td>
<td></td>
<td></td>
<td>B &gt; SE</td>
<td>NC</td>
</tr>
<tr>
<td><strong>Relative Level Change</strong></td>
<td></td>
<td></td>
<td>B &lt; SE</td>
<td>NC</td>
</tr>
<tr>
<td><strong>% Non Overlapping Data</strong></td>
<td></td>
<td></td>
<td>B = SE</td>
<td>B = PF</td>
</tr>
</tbody>
</table>

**Tallied Features**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>N Valid Features</th>
<th>Supported</th>
<th>Not Supported</th>
<th>Tied</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE &gt; Baseline</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>T</td>
</tr>
<tr>
<td>PF &gt; Baseline</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>T</td>
</tr>
<tr>
<td>PF &gt; SE</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>Y</td>
</tr>
</tbody>
</table>

*Note:* B = Baseline; PF = Performance Feedback; SE = Self-Evaluation; NC = Not Comparable Features; Y = Yes; N = No; T = Tied
Table 5

*Case D – Visual Analysis of Core Components*

<table>
<thead>
<tr>
<th></th>
<th>Stable/ Trend Dir</th>
<th>Baseline vs Self-Eval</th>
<th>Baseline vs PF</th>
<th>PF vs SE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong></td>
<td>Level Y (92%) IC</td>
<td>B &gt; PF</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trend Y (zero celerating) IC</td>
<td>B = PF</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Self-Evaluation (SE)</strong></td>
<td>Level N (83%) IC</td>
<td></td>
<td></td>
<td>IC</td>
</tr>
<tr>
<td></td>
<td>Trend N (decelerating) IC</td>
<td></td>
<td></td>
<td>IC</td>
</tr>
<tr>
<td><strong>Performance Feedback (PF)</strong></td>
<td>Level Y (91%)</td>
<td>B &gt; PF IC</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trend Y (accelerating)</td>
<td>B = PF IC</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Tallied Features**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>N Valid Features</th>
<th>Supported</th>
<th>Not Supported</th>
<th>Tied</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE &gt; Baseline</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IC</td>
</tr>
<tr>
<td>PF &gt; Baseline</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>T</td>
</tr>
<tr>
<td>PF &gt; SE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IC</td>
</tr>
</tbody>
</table>

*Note. B = Baseline; PF = Performance Feedback; SE = Self-Evaluation; IC = Invalid Comparison; NC = Not Comparable Features; Y = Yes; N = No; T = Tied*
Table 6

*Case E – Visual Analysis of Core Components*

<table>
<thead>
<tr>
<th></th>
<th>Stable/Trend Dir</th>
<th>Baseline vs Self-Eval</th>
<th>Baseline vs PF</th>
<th>PF vs SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>Level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Y (100%)</td>
<td>B = SE</td>
<td>B = PF</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Y (zero celerating)</td>
<td>B = SE</td>
<td>B = PF</td>
<td></td>
</tr>
<tr>
<td>Self-Evaluation (SE)</td>
<td>Level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Y (100%)</td>
<td>B = SE</td>
<td>PF = SE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Y (zero celerating)</td>
<td>B = SE</td>
<td>PF = SE</td>
<td></td>
</tr>
<tr>
<td>Performance Feedback (PF)</td>
<td>Level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Y (100%)</td>
<td>B = PF</td>
<td>PF = SE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Y (zero celerating)</td>
<td>B = PF</td>
<td>PF = SE</td>
<td></td>
</tr>
<tr>
<td>Absolute Level Change</td>
<td></td>
<td>B &lt; SE</td>
<td>NC</td>
<td>PF = SE</td>
</tr>
<tr>
<td>Relative Level Change</td>
<td></td>
<td>B = SE</td>
<td>NC</td>
<td>PF = SE</td>
</tr>
<tr>
<td>% Non Overlapping Data</td>
<td></td>
<td>B = SE</td>
<td>B = PF</td>
<td>PF = SE</td>
</tr>
</tbody>
</table>

**Tallied Features**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>N Valid Features</th>
<th>Supported</th>
<th>Not Supported</th>
<th>Tied</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE &gt; Baseline</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>T</td>
</tr>
<tr>
<td>PF &gt; Baseline</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>T</td>
</tr>
<tr>
<td>PF &gt; SE</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>T</td>
</tr>
</tbody>
</table>

*Note.* B = Baseline; PF = Performance Feedback; SE = Self-Evaluation; NC = Not Comparable Features; Y = Yes; N = No; T = Tied
Table 7

*Case F – Visual Analysis of Core Components*

<table>
<thead>
<tr>
<th>Stable/ Trend Dir</th>
<th>Baseline vs PF</th>
<th>Baseline vs SE</th>
<th>PF vs SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline Level</td>
<td>B &lt; PF</td>
<td>B &lt; SE</td>
<td></td>
</tr>
<tr>
<td>Trend</td>
<td>B = PF</td>
<td>B &gt; SE</td>
<td></td>
</tr>
<tr>
<td>Performance Feedback (PF)</td>
<td>Level (accelerating)</td>
<td>B &lt; PF</td>
<td>PF = SE</td>
</tr>
<tr>
<td>Trend</td>
<td>B = PF</td>
<td>PF &gt; SE</td>
<td></td>
</tr>
<tr>
<td>Self-Evaluation (SE)</td>
<td>Level (decelerating)</td>
<td>B &lt; SE</td>
<td>PF = SE</td>
</tr>
<tr>
<td>Trend</td>
<td>B &gt; SE</td>
<td>PF &gt; SE</td>
<td></td>
</tr>
<tr>
<td>Absolute Level Change</td>
<td></td>
<td>B &lt; PF</td>
<td>NC</td>
</tr>
<tr>
<td>Relative Level Change</td>
<td></td>
<td>B &lt; PF</td>
<td>NC</td>
</tr>
<tr>
<td>% Non Overlapping Data</td>
<td></td>
<td>B &lt; PF</td>
<td>B = SE</td>
</tr>
</tbody>
</table>

**Tallied Features**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>N Valid Features</th>
<th>Supported</th>
<th>Not Supported</th>
<th>Tied</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>PF &gt; Baseline</td>
<td>5</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>Y</td>
</tr>
<tr>
<td>SE &gt; Baseline</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>T</td>
</tr>
<tr>
<td>PF &gt; SE</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>T</td>
</tr>
</tbody>
</table>

*Note. B = Baseline; PF = Performance Feedback; SE = Self-Evaluation; NC = Not Comparable Features; Y = Yes; N = No; T = Tied*
Table 8

*Case G – Visual Analysis of Core Components*

<table>
<thead>
<tr>
<th></th>
<th>Stable/ Trend Dir</th>
<th>Baseline vs PF</th>
<th>Baseline vs SE</th>
<th>PF vs SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>Level</td>
<td>B &lt; PF</td>
<td>B &lt; SE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>B &lt; PF</td>
<td>B &lt; SE</td>
<td></td>
</tr>
<tr>
<td>Performance Feedback (PF)</td>
<td>Level</td>
<td>B &lt; PF</td>
<td>PF &lt; SE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>B &lt; PF</td>
<td>PF = SE</td>
<td></td>
</tr>
<tr>
<td>Self-Evaluation (SE)</td>
<td>Level</td>
<td>B &lt; SE</td>
<td>PF &lt; SE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>B &lt; SE</td>
<td>PF = SE</td>
<td></td>
</tr>
<tr>
<td>Absolute Level Change</td>
<td></td>
<td>B &lt; PF</td>
<td>NC</td>
<td>PF &lt; SE</td>
</tr>
<tr>
<td>Relative Level Change</td>
<td></td>
<td>B &lt; PF</td>
<td>NC</td>
<td>PF = SE</td>
</tr>
<tr>
<td>% Non Overlapping Data</td>
<td></td>
<td>B = PF</td>
<td>B &lt; SE</td>
<td>PF = SE</td>
</tr>
</tbody>
</table>

**Tallied Features**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>N Valid Features</th>
<th>Supported</th>
<th>Not Supported</th>
<th>Tied</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>PF &gt; Baseline</td>
<td>5</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>Y</td>
</tr>
<tr>
<td>SE &gt; Baseline</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>Y</td>
</tr>
<tr>
<td>PF &gt; SE</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>T</td>
</tr>
</tbody>
</table>

*Note:* B = Baseline; PF = Performance Feedback; SE = Self-Evaluation; NC = Not Comparable Features; Y = Yes; N = No; T = Tied
Table 9

**Case H – Visual Analysis of Core Components**

<table>
<thead>
<tr>
<th></th>
<th>Stable/Trend Dir</th>
<th>Baseline vs PF</th>
<th>Baseline vs SE</th>
<th>PF vs SE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level</td>
<td>Y (75%)</td>
<td>IC</td>
<td>B &lt; SE</td>
<td></td>
</tr>
<tr>
<td>Trend</td>
<td>Y (accelerating)</td>
<td>IC</td>
<td>B = SE</td>
<td></td>
</tr>
<tr>
<td><strong>Performance Feedback (PF)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level</td>
<td>N (91%)</td>
<td>IC</td>
<td>IC</td>
<td></td>
</tr>
<tr>
<td>Trend</td>
<td>N (accelerating)</td>
<td>IC</td>
<td>IC</td>
<td></td>
</tr>
<tr>
<td><strong>Self-Evaluation (SE)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level</td>
<td>Y (96%)</td>
<td>B &lt; SE</td>
<td>IC</td>
<td></td>
</tr>
<tr>
<td>Trend</td>
<td>Y (accelerating)</td>
<td>B = SE</td>
<td>IC</td>
<td></td>
</tr>
<tr>
<td><strong>Absolute Level Change</strong></td>
<td></td>
<td>IC</td>
<td>NC</td>
<td>IC</td>
</tr>
<tr>
<td><strong>Relative Level Change</strong></td>
<td></td>
<td>IC</td>
<td>NC</td>
<td>IC</td>
</tr>
<tr>
<td><strong>% Non Overlapping Data</strong></td>
<td></td>
<td>IC</td>
<td>B &lt; SE</td>
<td>IC</td>
</tr>
</tbody>
</table>

**Tallied Features**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>N Valid Features</th>
<th>Supported</th>
<th>Not Supported</th>
<th>Tied</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>PF &gt; Baseline</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>Y</td>
</tr>
<tr>
<td>SE &gt; Baseline</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PF &gt; SE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IC</td>
</tr>
</tbody>
</table>

*Note. B = Baseline; PF = Performance Feedback; SE = Self-Evaluation; IC = Invalid Comparison; NC = Not Comparable Features; Y = Yes; N = No; T = Tied*
Table 10

*Case I – Visual Analysis of Core Components*

<table>
<thead>
<tr>
<th></th>
<th>Stable/ Trend Dir</th>
<th>Baseline vs PF</th>
<th>Baseline vs SE</th>
<th>PF vs SE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong></td>
<td>Level</td>
<td>B &lt; PF</td>
<td>B &lt; SE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>B = PF</td>
<td>B = SE</td>
<td></td>
</tr>
<tr>
<td><strong>Performance Feedback (PF)</strong></td>
<td>Level</td>
<td>B &lt; PF</td>
<td>PF = SE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>B = PF</td>
<td>PF = SE</td>
<td></td>
</tr>
<tr>
<td><strong>Self-Evaluation (SE)</strong></td>
<td>Level</td>
<td>B &lt; SE</td>
<td>PF = SE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>B = SE</td>
<td>PF = SE</td>
<td></td>
</tr>
<tr>
<td><strong>Absolute Level Change</strong></td>
<td></td>
<td>B &gt; PF</td>
<td>NC</td>
<td>PF = SE</td>
</tr>
<tr>
<td><strong>Relative Level Change</strong></td>
<td></td>
<td>B &gt; PF</td>
<td>NC</td>
<td>PF = SE</td>
</tr>
<tr>
<td><strong>% Non Overlapping Data</strong></td>
<td></td>
<td>B = PF</td>
<td>B = SE</td>
<td>PF = SE</td>
</tr>
</tbody>
</table>

**Tallied Features**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>N Valid Features</th>
<th>Supported</th>
<th>Not Supported</th>
<th>Tied</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>PF &gt; Baseline</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>T</td>
</tr>
<tr>
<td>SE &gt; Baseline</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>T</td>
</tr>
<tr>
<td>PF &gt; SE</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>T</td>
</tr>
</tbody>
</table>

*Note:* B = Baseline; PF = Performance Feedback; SE = Self-Evaluation; NC = Not Comparable Features; Y = Yes; N = No; T = Tied
Visual Analysis of the HELPS Tips and Reminders Treatment Integrity

The percentage of tips and reminders correctly implemented each session by parent participants is presented in Figures 3 and 4. Data points that are not filled represent sessions in which I directly observed the parent implementing the HELPS Program, and depending on the phase, where I instructed the parent to complete the self-evaluation procedures or I provided performance feedback.

Parents in cases A through E, where self-evaluation was the first feedback method, demonstrated varying integrity levels during the baseline phase. The median integrity level of each phase for each case is reported in Table 11.

Table 11

<table>
<thead>
<tr>
<th>ABC Cohort</th>
<th>Baseline</th>
<th>Self-Evaluation</th>
<th>Performance Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case A</td>
<td>41.5%</td>
<td>60%</td>
<td>69%</td>
</tr>
<tr>
<td></td>
<td>(39-45)</td>
<td>(36-83)</td>
<td>(52-82)</td>
</tr>
<tr>
<td>Case B</td>
<td>24%</td>
<td>33%</td>
<td>25%</td>
</tr>
<tr>
<td></td>
<td>(14-36)</td>
<td>(26-52)</td>
<td>(17-58)</td>
</tr>
<tr>
<td>Case C</td>
<td>65%</td>
<td>74%</td>
<td>88%</td>
</tr>
<tr>
<td></td>
<td>(57-82)</td>
<td>(65-78)</td>
<td>(70-96)</td>
</tr>
<tr>
<td>Case D</td>
<td>69%</td>
<td>74%</td>
<td>85%</td>
</tr>
<tr>
<td></td>
<td>(67-85)</td>
<td>(57-87)</td>
<td>(57-96)</td>
</tr>
<tr>
<td>Case E</td>
<td>75.5%</td>
<td>79%</td>
<td>94%</td>
</tr>
<tr>
<td></td>
<td>(63-92)</td>
<td>(67-95)</td>
<td>(82-96)</td>
</tr>
</tbody>
</table>
Two cases (Case A & E) had treatment integrity that declined during the baseline phase. With the introduction of self-evaluation the treatment integrity immediately improved for both cases; however, for Case E the dramatic level increase from 63% to 85% began declining after three sessions. For Case A the treatment integrity improved over time. Cases B and C had improving treatment integrity during the baseline phase that continued to improve during the self-evaluation phase. Case D had a zero celerating trend during the baseline phase. The introduction of self-evaluation shows some improvement in treatment integrity for some of the sessions, however, the data trend reflects declining treatment integrity.

When Cases A through E entered the performance feedback phase, the data began improving across time, as demonstrated by an accelerating trend line across all cases. For Cases A, B, and C, the trend line in the preceding phase, the self-evaluation phase, was
Figure 3. Treatment integrity of HELPS Program tips and reminders for participants in ABC cohort. ◆ = session not directly observed; ◇ = direct observation/feedback for Phases 2 & 3
Figure 4. Treatment integrity of HELPS Program tips and reminders for participants in ACB cohort. ◆ = session not directly observed; ◇ = direct observation/feedback for Phases 2 & 3
accelerating as well. For Cases D and E, treatment integrity in the performance feedback phase improved over time, as compared to the self-evaluation phase where treatment integrity declined over time. Case D data for the performance feedback phase, however, was unstable. For Case E, there was an immediate level change with the introduction of performance feedback, with the level improving from 67% to 96%.

The second cohort, Cases F through I, received performance feedback as the first feedback phase. Three of the four cases (Cases F, G, & I) had baseline integrity data that declined over time, as demonstrated by a stable decelerating trend line. With the introduction of performance feedback, there was an immediate level change for all three cases, demonstrating improved treatment integrity. All three cases had data in the performance feedback phase that reflected a stable accelerating trend line. Case H had baseline treatment integrity data that stable and showing a general trend toward improvement over time. Although there was an immediate decline in treatment integrity following the introduction of performance feedback, the stable trend line demonstrated improvement over time.

When Cases F through I entered the self-evaluation phase the data across cases reflect less consistency. In Cases F and H the data trend changed from accelerating in the performance feedback phase to decelerating in the self-evaluation phase, indicating a decline in treatment integrity. For Cases G and I the treatment integrity continued to improve following the introduction of self-evaluation.

Summary of the visual analysis of the tips and reminders. Across cases, five of the nine cases (Cases A, E, F, G, & I) exhibited declining baseline data. In four of the five cases (Cases A, F, G, & I) the introduction of a feedback method reflected an immediate improved
level change in the treatment integrity, followed by general improvement in integrity across the first feedback phase. The baseline data for Case E reflected a decelerating baseline trend, and although there was an immediate improved level change with the introduction of feedback, the data trend reflected a decline over the length of the phase. Additionally, four of the five cases (Cases A, E, G, & I) demonstrated higher treatment integrity level in the feedback phases than in the baseline phase. Of the remaining four cases (Cases B, C, D, & H) all cases but one (Case D) had data in the baseline and first feedback phase that was stable and reflected stable acceleration over time. Case D had stable zero celerating baseline data and stable decelerating data in the first feedback phase. Thus, in the majority of cases, following baseline, the introduction of a feedback method reflected improved treatment integrity, regardless of what form of feedback was provided.

In four of the nine cases (Cases A, C, G, & I) treatment integrity in both feedback phases reflected stable accelerating trend lines. Case B also had accelerating data in both feedback phases; however, the data were unstable in the second feedback phase. Generally, however, in the majority of cases treatment integrity continued to improve over time, regardless of which feedback method was provided first. In the remaining cases where data in both feedback phases was not accelerating (Cases D, E, F, & H), all four cases show greater treatment integrity in the performance feedback phase, regardless of the order in which performance feedback was provided (either before or after self-evaluation).

Lastly, when comparing median integrity level across all cases, there is a general trend of improvement over time across phases. Again, regardless of sequence of feedback methods, participants generally demonstrated improved treatment integrity level as the
sessions progressed.

The results of the visual analysis of the tips and reminders for each case are summarized in Tables 12 through 20.
Table 12

Case A – Visual Analysis of Tips and Reminders

<table>
<thead>
<tr>
<th>Stable/Trend Dir</th>
<th>Baseline vs Self-Eval</th>
<th>Baseline vs PF</th>
<th>PF vs SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>Level</td>
<td>IC</td>
<td>IC</td>
</tr>
<tr>
<td></td>
<td>(41.5%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>IC</td>
<td>B &lt; PF</td>
</tr>
<tr>
<td></td>
<td>(decelerating)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-Evaluation</td>
<td>Level</td>
<td>IC</td>
<td>IC</td>
</tr>
<tr>
<td>(SE)</td>
<td>(60%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>IC</td>
<td>IC</td>
</tr>
<tr>
<td></td>
<td>(accelerating)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance</td>
<td>Level</td>
<td>IC</td>
<td>IC</td>
</tr>
<tr>
<td>Feedback (PF)</td>
<td>(69%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>B &lt; PF</td>
<td>IC</td>
</tr>
<tr>
<td></td>
<td>(accelerating)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absolute Level</td>
<td>IC</td>
<td>NC</td>
<td>IC</td>
</tr>
<tr>
<td>Change</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative Level</td>
<td>IC</td>
<td>NC</td>
<td>IC</td>
</tr>
<tr>
<td>Change</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Non Overlapping</td>
<td>IC</td>
<td>B &lt; PF</td>
<td>IC</td>
</tr>
<tr>
<td>Data</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Tallied Features

<table>
<thead>
<tr>
<th>Comparison</th>
<th>N Valid Features</th>
<th>Supported</th>
<th>Not Supported</th>
<th>Tied</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE &gt; Baseline</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>IC</td>
</tr>
<tr>
<td>PF &gt; Baseline</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>Y</td>
</tr>
<tr>
<td>PF &gt; SE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IC</td>
</tr>
</tbody>
</table>

Note. B = Baseline; PF = Performance Feedback; SE = Self-Evaluation; IC = Invalid Comparison; NC = Not Comparable Features; Y = Yes; N = No; T = Tied
Table 13

*Case B – Visual Analysis of Tips and Reminders*

<table>
<thead>
<tr>
<th></th>
<th>Stable/ Trend Dir</th>
<th>Baseline vs Self-Eval</th>
<th>Baseline vs PF</th>
<th>PF vs SE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong></td>
<td>Level</td>
<td>B &lt; SE</td>
<td>IC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>B = SE</td>
<td>IC</td>
<td></td>
</tr>
<tr>
<td><strong>Self-Evaluation (SE)</strong></td>
<td>Level</td>
<td>B &lt; SE</td>
<td>IC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>B = SE</td>
<td>IC</td>
<td></td>
</tr>
<tr>
<td><strong>Performance Feedback (PF)</strong></td>
<td>Level</td>
<td>B = SE</td>
<td>IC</td>
<td>IC</td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>B = SE</td>
<td>IC</td>
<td>IC</td>
</tr>
</tbody>
</table>

- **Absolute Level Change**
  - B < SE
  - NC
  - IC

- **Relative Level Change**
  - B < SE
  - NC
  - IC

- **% Non Overlapping Data**
  - B = SE
  - IC
  - IC

**Tallied Features**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>N Valid Features</th>
<th>Supported</th>
<th>Not Supported</th>
<th>Tied</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE &gt; Baseline</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>Y</td>
</tr>
<tr>
<td>PF &gt; Baseline</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IC</td>
</tr>
<tr>
<td>PF &gt; SE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IC</td>
</tr>
</tbody>
</table>

*Note.* B = Baseline; PF = Performance Feedback; SE = Self-Evaluation; IC = Invalid Comparison; NC = Not Comparable Features; Y = Yes; N = No; T = Tied
Table 14

*Case C – Visual Analysis of Tips and Reminders*

<table>
<thead>
<tr>
<th></th>
<th>Stable/Trend Dir</th>
<th>Baseline vs Self-Eval</th>
<th>Baseline vs PF</th>
<th>PF vs SE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong></td>
<td>Level</td>
<td>Y (65%)</td>
<td>B &lt; SE</td>
<td>B &lt; PF</td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>Y (accelerating)</td>
<td>B = SE</td>
<td>B = PF</td>
</tr>
<tr>
<td><strong>Self-Evaluation (SE)</strong></td>
<td>Level</td>
<td>Y (74%)</td>
<td>B &lt; SE</td>
<td>PF &gt; SE</td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>Y (accelerating)</td>
<td>B = SE</td>
<td>PF = SE</td>
</tr>
<tr>
<td><strong>Performance Feedback (PF)</strong></td>
<td>Level</td>
<td>Y (88%)</td>
<td>B &lt; PF</td>
<td>PF &gt; SE</td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>Y (accelerating)</td>
<td>B = PF</td>
<td>PF = SE</td>
</tr>
<tr>
<td><strong>Absolute Level Change</strong></td>
<td></td>
<td></td>
<td>B &gt; SE</td>
<td>NC</td>
</tr>
<tr>
<td><strong>Relative Level Change</strong></td>
<td></td>
<td></td>
<td>B &lt; SE</td>
<td>NC</td>
</tr>
<tr>
<td><strong>% Non Overlapping Data</strong></td>
<td></td>
<td></td>
<td>B = SE</td>
<td>B &lt; PF</td>
</tr>
</tbody>
</table>

**Tallied Features**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>N Valid Features</th>
<th>Valid Supported</th>
<th>Not Supported</th>
<th>Tied</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE &gt; Baseline</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>T</td>
</tr>
<tr>
<td>PF &gt; Baseline</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>Y</td>
</tr>
<tr>
<td>PF &gt; SE</td>
<td>5</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>Y</td>
</tr>
</tbody>
</table>

*Note.* B = Baseline; PF = Performance Feedback; SE = Self-Evaluation; IC = Invalid Comparison; NC = Not Comparable Features; Y = Yes; N = No; T = Tied
Table 15

*Case D – Visual Analysis of Tips and Reminders*

<table>
<thead>
<tr>
<th></th>
<th>Stable/ Trend Dir</th>
<th>Baseline vs Self-Eval</th>
<th>Baseline vs PF</th>
<th>PF vs SE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong></td>
<td>Level</td>
<td>Y (69%)</td>
<td>IC</td>
<td>IC</td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>Y (zero accelerating)</td>
<td>B = SE</td>
<td>IC</td>
</tr>
<tr>
<td><strong>Self-Evaluation (SE)</strong></td>
<td>Level</td>
<td>N (74%)</td>
<td>IC</td>
<td>IC</td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>Y (decelerating)</td>
<td>B = SE</td>
<td>IC</td>
</tr>
<tr>
<td><strong>Performance Feedback (PF)</strong></td>
<td>Level</td>
<td>N (85%)</td>
<td>IC</td>
<td>IC</td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>N (accelerating)</td>
<td>IC</td>
<td>IC</td>
</tr>
<tr>
<td><strong>Absolute Level Change</strong></td>
<td></td>
<td>B &gt; SE</td>
<td>NC</td>
<td>IC</td>
</tr>
<tr>
<td><strong>Relative Level Change</strong></td>
<td></td>
<td>B &lt; SE</td>
<td>NC</td>
<td>IC</td>
</tr>
<tr>
<td><strong>% Non Overlapping Data</strong></td>
<td></td>
<td>B = SE</td>
<td>IC</td>
<td>IC</td>
</tr>
</tbody>
</table>

**Tallied Features**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>N Valid Features</th>
<th>Supported</th>
<th>Not Supported</th>
<th>Tied</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE &gt; Baseline</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>T</td>
</tr>
<tr>
<td>PF &gt; Baseline</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IC</td>
</tr>
<tr>
<td>PF &gt; SE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IC</td>
</tr>
</tbody>
</table>

*Note. B = Baseline; PF = Performance Feedback; SE = Self-Evaluation; IC = Invalid Comparison; NC = Not Comparable Features; Y = Yes; N = No; T = Tied*
Table 16

Case E – Visual Analysis of Tips and Reminders

<table>
<thead>
<tr>
<th></th>
<th>Stable/Trend Dir</th>
<th>Baseline vs Self-Eval</th>
<th>Baseline vs PF</th>
<th>PF vs SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>Level</td>
<td>Y</td>
<td>B &lt; SE</td>
<td>B &lt; PF</td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>Y (75.5%)</td>
<td>B = SE</td>
<td>B &lt; PF</td>
</tr>
<tr>
<td>Self-Evaluation (SE)</td>
<td>Level</td>
<td>Y</td>
<td>B &lt; SE</td>
<td>PF &gt; SE</td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>Y (decelerating)</td>
<td>B = SE</td>
<td>PF &gt; SE</td>
</tr>
<tr>
<td>Performance Feedback (PF)</td>
<td>Level</td>
<td>Y</td>
<td>B &lt; PF</td>
<td>PF &gt; SE</td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>Y (accelerating)</td>
<td>B &lt; PF</td>
<td>PF &gt; SE</td>
</tr>
<tr>
<td>Absolute Level Change</td>
<td></td>
<td></td>
<td>B &lt; SE</td>
<td>NC</td>
</tr>
<tr>
<td>Relative Level Change</td>
<td></td>
<td></td>
<td>B &lt; SE</td>
<td>NC</td>
</tr>
<tr>
<td>% Non Overlapping Data</td>
<td></td>
<td></td>
<td>B = SE</td>
<td>B = PF</td>
</tr>
</tbody>
</table>

Tallied Features

<table>
<thead>
<tr>
<th>Comparison</th>
<th>N Valid Features</th>
<th>Supported</th>
<th>Not Supported</th>
<th>Tied</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>SE &gt; Baseline</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>Y</td>
</tr>
<tr>
<td>PF &gt; Baseline</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>Y</td>
</tr>
<tr>
<td>PF &gt; SE</td>
<td>5</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>Y</td>
</tr>
</tbody>
</table>

*Note:* B = Baseline; PF = Performance Feedback; SE = Self-Evaluation; NC = Not Comparable Features; Y = Yes; N = No; T = Tied
Table 17

*Case F – Visual Analysis of Tips and Reminders*

<table>
<thead>
<tr>
<th>Stable/Trend Dir</th>
<th>Baseline vs PF</th>
<th>Baseline vs SE</th>
<th>PF vs SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>Level</td>
<td>B &lt; PF</td>
<td>IC</td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>B &lt; PF</td>
<td>B = SE</td>
</tr>
<tr>
<td>Performance Feedback (PF)</td>
<td>Level</td>
<td>B &lt; PF</td>
<td>IC</td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>B &lt; PF</td>
<td>PF &gt; SE</td>
</tr>
<tr>
<td>Self-Evaluation (SE)</td>
<td>Level</td>
<td>IC</td>
<td>IC</td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>B = SE</td>
<td>PF &gt; SE</td>
</tr>
<tr>
<td>Absolute Level Change</td>
<td></td>
<td>B &lt; PF</td>
<td>NC</td>
</tr>
<tr>
<td>Relative Level Change</td>
<td></td>
<td>B &lt; PF</td>
<td>NC</td>
</tr>
<tr>
<td>% Non Overlapping Data</td>
<td></td>
<td>B &lt; PF</td>
<td>B = SE</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tallied Features</th>
<th>N Valid Features</th>
<th>Supported</th>
<th>Not Supported</th>
<th>Tied</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>PF &gt; Baseline</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>Y</td>
</tr>
<tr>
<td>SE &gt; Baseline</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>T</td>
</tr>
<tr>
<td>PF &gt; SE</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>Y</td>
</tr>
</tbody>
</table>

*Note.* B = Baseline; PF = Performance Feedback; SE = Self-Evaluation; IC = Invalid Comparison; NC = Not Comparable Features; Y = Yes; N = No; T = Tied
Table 18

Case G – Visual Analysis of Tips and Reminders

<table>
<thead>
<tr>
<th></th>
<th>Stable/Trend Dir</th>
<th>Baseline vs PF</th>
<th>Baseline vs SE</th>
<th>PF vs SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>Level</td>
<td>B &lt; PF</td>
<td>B &lt; SE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>B &lt; PF</td>
<td>B &lt; SE</td>
<td></td>
</tr>
<tr>
<td>Performance Feedback (PF)</td>
<td>Level</td>
<td>B &lt; PF</td>
<td>PF &lt; SE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>B &lt; PF</td>
<td>PF = SE</td>
<td></td>
</tr>
<tr>
<td>Self-Evaluation (SE)</td>
<td>Level</td>
<td>B &lt; SE</td>
<td>PF &lt; SE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>B &lt; SE</td>
<td>PF = SE</td>
<td></td>
</tr>
</tbody>
</table>

Absolute Level Change
Relative Level Change
% Non Overlapping Data

<table>
<thead>
<tr>
<th></th>
<th>Stable/Trend Dir</th>
<th>Baseline vs PF</th>
<th>Baseline vs SE</th>
<th>PF vs SE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B &lt; PF</td>
<td>NC</td>
<td>PF &gt; SE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B &lt; PF</td>
<td>NC</td>
<td>PF &lt; SE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>B &lt; PF</td>
<td>B &lt; SE</td>
<td>PF = SE</td>
<td></td>
</tr>
</tbody>
</table>

Tallied Features

<table>
<thead>
<tr>
<th>Comparison</th>
<th>N Valid Features</th>
<th>Supported</th>
<th>Not Supported</th>
<th>Tied</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>PF &gt; Baseline</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>Y</td>
</tr>
<tr>
<td>SE &gt; Baseline</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>Y</td>
</tr>
<tr>
<td>PF &gt; SE</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>T</td>
</tr>
</tbody>
</table>

*Note.* B = Baseline; PF = Performance Feedback; SE = Self-Evaluation; NC = Not Comparable Features; Y = Yes; N = No; T = Tied
Table 19

*Case H – Visual Analysis of Tips and Reminders*

<table>
<thead>
<tr>
<th></th>
<th>Stable/Trend Dir</th>
<th>Baseline vs PF</th>
<th>Baseline vs SE</th>
<th>PF vs SE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong> Level Y (62%)</td>
<td>IC</td>
<td>B &lt; SE</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Baseline</strong> Trend Y</td>
<td>B = PF</td>
<td>B &gt; SE</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Performance Feedback (PF)</strong></td>
<td>Level N (73%)</td>
<td>IC</td>
<td>IC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trend Y</td>
<td>B = PF</td>
<td>PF &gt; SE</td>
<td></td>
</tr>
<tr>
<td><strong>Self-Evaluation (SE)</strong></td>
<td>Level Y (74%)</td>
<td>B &lt; SE</td>
<td>IC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trend Y</td>
<td>B &gt; SE</td>
<td>PF &gt; SE</td>
<td></td>
</tr>
<tr>
<td><strong>Absolute Level Change</strong></td>
<td></td>
<td>B &gt; PF</td>
<td>NC</td>
<td>PF &lt; SE</td>
</tr>
<tr>
<td><strong>Relative Level Change</strong></td>
<td></td>
<td>B &gt; PF</td>
<td>NC</td>
<td>PF &gt; SE</td>
</tr>
<tr>
<td><strong>% Non Overlapping Data</strong></td>
<td></td>
<td>B = PF</td>
<td>B = SE</td>
<td>PF = SE</td>
</tr>
</tbody>
</table>

**Tallied Features**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>N Valid Features</th>
<th>Supported</th>
<th>Not Supported</th>
<th>Tied</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>PF &gt; Baseline</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>T</td>
</tr>
<tr>
<td>SE &gt; Baseline</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>T</td>
</tr>
<tr>
<td>PF &gt; SE</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>T</td>
</tr>
</tbody>
</table>

*Note.* B = Baseline; PF = Performance Feedback; SE = Self-Evaluation; IC = Invalid Comparison; NC = Not Comparable Features; Y = Yes; N = No; T = Tied
### Table 20

*Case I – Visual Analysis of Tips and Reminders*

<table>
<thead>
<tr>
<th></th>
<th>Stable/Trend Dir</th>
<th>Baseline vs PF</th>
<th>Baseline vs SE</th>
<th>PF vs SE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong></td>
<td>Level</td>
<td>B &lt; PF</td>
<td>B &lt; SE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>B &lt; PF</td>
<td>B &lt; SE</td>
<td></td>
</tr>
<tr>
<td><strong>Performance Feedback (PF)</strong></td>
<td>Level</td>
<td>B &lt; PF</td>
<td>PF &lt; SE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>B &lt; PF</td>
<td>PF = SE</td>
<td></td>
</tr>
<tr>
<td><strong>Self-Evaluation (SE)</strong></td>
<td>Level</td>
<td>B &lt; SE</td>
<td>PF &lt; SE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trend</td>
<td>B &lt; SE</td>
<td>PF = SE</td>
<td></td>
</tr>
<tr>
<td><strong>Absolute Level Change</strong></td>
<td></td>
<td>B &lt; PF</td>
<td>NC</td>
<td>PF &gt; SE</td>
</tr>
<tr>
<td><strong>Relative Level Change</strong></td>
<td></td>
<td>B &lt; PF</td>
<td>NC</td>
<td>PF = SE</td>
</tr>
<tr>
<td><strong>% Non Overlapping Data</strong></td>
<td></td>
<td>B &lt; PF</td>
<td>B &lt; SE</td>
<td>PF = SE</td>
</tr>
</tbody>
</table>

#### Tallied Features

<table>
<thead>
<tr>
<th>Comparison</th>
<th>N Valid Features</th>
<th>Supported</th>
<th>Not Supported</th>
<th>Tied</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>PF &gt; Baseline</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>Y</td>
</tr>
<tr>
<td>SE &gt; Baseline</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>Y</td>
</tr>
<tr>
<td>PF &gt; SE</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>T</td>
</tr>
</tbody>
</table>

*Note.* B = Baseline; PF = Performance Feedback; SE = Self-Evaluation; NC = Not Comparable Features; Y = Yes; N = No; T = Tied
To answer my research questions and hypotheses using a sign test (Corder & Foreman, 2014), I needed at least 5 cases that showed a difference between phases favoring one condition over the other to detect a reliable difference at the p < .05 threshold.

**Research questions**

**Research question 1:** Is treatment integrity of the HELPS intervention core components higher when parents receive performance feedback than when parents receive self-evaluative feedback?

Of the 9 cases reflecting treatment integrity of the core components, 2 cases (Cases A & C) showed a difference in treatment integrity between the performance feedback phase and the self-evaluation phase, with both cases demonstrating higher treatment integrity in the performance feedback phase. The remaining 7 cases either did not have a level and/or trend that was sufficiently stable to compare (Cases D & H), and were therefore excluded from the analysis, or the comparison of the two phases did not result in a clear difference between the phases (Cases B, E, F, G, & I). The 5 cases showing no clear difference between phases were therefore discarded from the analysis. Due to the reduction in the number of cases that showed a difference in phases (N = 2), I did not have enough cases to conduct a sign test (Corder & Foreman) and detect an effect at the p < .05 threshold.

**Research question 2:** Is treatment integrity of the HELPS intervention tips and reminders higher when parents receive performance feedback than when parents receive self-evaluative feedback?

Of the 9 cases reflecting treatment integrity of the tips and reminders, 3 cases (Cases C, E, & F) showed a difference in treatment integrity between the performance feedback
phase and the self-evaluation phase, with all 3 cases demonstrating higher treatment integrity during the performance feedback phase. The remaining 6 cases either did not have a level and/or trend that was sufficiently stable to compare (Case A, B, & D), and were therefore excluded from the analysis, or the comparison of the two phases did not result in a clear difference between the phases (Cases G, H, & I). The 3 cases showing no clear difference between phases were therefore discarded from the analysis. Due to the reduction in the number of cases that showed a difference in phases ($N = 3$), I did not have enough cases to conduct a sign test (Corder & Foreman, 2014) and detect an effect at the $p < .05$ threshold.

**Hypotheses**

**Hypothesis 1:** Treatment integrity of the HELPS intervention core components will be higher in the performance feedback condition than in the baseline condition.

Of the 9 cases reflecting treatment integrity of the core components during the baseline and performance feedback phases, 2 cases (Cases F & G) showed a difference in treatment integrity between the baseline phase and the performance feedback phase, with both cases demonstrating higher treatment integrity during the performance feedback phase. The remaining 7 cases either did not have a level and/or trend that was sufficiently stable to compare (Case H), and was therefore excluded from the analysis, or the comparison of the two phases did not result in a clear difference between the phases (Cases A, B, C, D, E, & I). The 6 cases showing no clear difference between phases were therefore discarded from the analysis. Due to the reduction in the number of cases that showed a difference in phases ($N = 2$), I did not have enough cases to conduct a sign test (Corder & Foreman) and detect an effect at the $p < .05$ threshold.
**Hypothesis 2:** Treatment integrity of the HELPS intervention tips and reminders will be higher in the performance feedback condition than in the baseline condition.

Of the 9 cases reflecting treatment integrity of the tips and reminders during the baseline and performance feedback phases, 6 cases (Cases A, C, E, F, G, & I) showed a difference in treatment integrity between the baseline phase and the performance feedback phase, with all 6 cases demonstrating higher treatment integrity during the performance feedback phase. The remaining 3 cases either did not have a level and/or trend that was sufficiently stable to compare (Cases B & D), and were therefore excluded from the analysis, or the comparison of the two phases did not result in a clear difference between the phases (Case H). Case H showed no clear difference between phases, and was therefore discarded from the analysis. I conducted a Sign Test (Corder and Foreman) on the 6 cases demonstrating a difference between the baseline and performance feedback phases. The obtained value, $p = .0156$, was less than the critical value, $\alpha = 0.05$. Therefore, my hypothesis was supported.

**Hypothesis 3:** Treatment integrity of the HELPS intervention core components will be higher in the self-evaluation feedback condition than in the baseline condition.

Of the 9 cases reflecting treatment integrity of the core components during the baseline and self-evaluation phases, 4 cases (Cases A, B, G, & H) showed a difference in treatment integrity between the baseline phase and the self-evaluations phases, with all 4 cases demonstrating higher treatment integrity during the self-evaluation phase. The remaining 5 cases either did not have a level and/or trend that was sufficiently stable to compare (Case D), and was therefore excluded from the analysis, or the comparison of the
two phases did not result in a clear difference between the phases (Cases C, E, F, & I). The 4 cases showing no clear difference between phases were therefore discarded from the analysis. Due to the reduction in the number of cases that showed a difference in phases ($N = 4$), I did not have enough cases to conduct a sign test (Corder & Foreman) and detect an effect at the $p < .05$ threshold.

**Hypothesis 4:** Treatment integrity of the HELPS intervention tips and reminders will be higher in the self-evaluation feedback condition than in the baseline condition.

Of the 9 cases reflecting treatment integrity of the tips and reminders during the baseline and self-evaluation phases, 4 cases (Cases B, E, G, & I) showed a difference in treatment integrity between the baseline phase and the self-evaluations phases, with all 4 cases demonstrating higher treatment integrity during the self-evaluation phase. The remaining 5 cases either did not have a level and/or trend that was sufficiently stable to compare (Case A), and was therefore excluded from the analysis, or the comparison of the two phases did not result in a clear difference between the phases (Cases C, D, F, & H). The 4 cases showing no clear difference between phases were therefore discarded from the analysis. Due to the reduction in the number of cases that showed a difference in phases ($N = 4$), I did not have enough cases to conduct a sign test (Corder & Foreman) and detect an effect at the $p < .05$ threshold.

**Supplementary Analyses**

To further analyze the possible change in treatment integrity from the baseline phase to a feedback method, I compared each participant’s baseline treatment integrity level for the core components with the treatment integrity level for the core components during the first
feedback method, regardless of whether it was performance feedback or self-evaluation. Of the 9 cases reflecting treatment integrity of the core components during the baseline and first feedback phase, 4 cases (Cases A [SE], B [SE], F [PF], & G [PF]) showed a difference in treatment integrity between the baseline phase and the first feedback phase, with all 4 cases demonstrating higher treatment integrity during the feedback phase. The remaining 5 cases either did not have a level and/or trend that was sufficiently stable to compare (Cases D & H), and were therefore excluded from the analysis, or the comparison of the two phases did not result in a clear difference between the phases (Cases C, E, & I). The 3 cases showing no clear difference between phases were therefore discarded from the analysis. Due to the reduction in the number of cases that showed a difference in phases ($N = 4$), I did not have enough cases to conduct a sign test (Corder & Foreman) and detect an effect at the $p < .05$ threshold.

I also compared each participant’s baseline treatment integrity level for the tips and reminders with the treatment integrity level for the tips and reminders during the first feedback method, regardless of whether it was performance feedback or self-evaluation. Of the 9 cases reflecting treatment integrity of the tips and reminders during the baseline and first feedback phase, 5 cases (Cases B [SE], E [SE], F [PF], G [PF], & I [PF]) showed a difference in treatment integrity between the baseline phase and the first feedback phase, with all 5 cases demonstrating higher treatment integrity of the tips and reminders during the feedback phase than during the baseline phase. The remaining 4 cases either did not have a level and/or trend that was sufficiently stable to compare (Case A), and was therefore excluded from the analysis, or the comparison of the two phases did not result in a clear
difference between the phases (Cases C, D, & H). The 3 cases showing no clear difference between phases were therefore discarded from the analysis. I conducted a Sign Test (Corder and Foreman) on the 5 cases demonstrating a difference between the baseline and first feedback phase. The obtained value, \( p = .0313 \), was less than the critical value, \( \alpha = 0.05 \), suggesting that there is a reliable difference in the treatment integrity of the tips and reminders between the baseline and first feedback phase.

Based on the visual analysis I examined each case for a general trend in treatment integrity improvement of the core components and the tips and reminders over time, from baseline through phase three. By examining each case based on level and trend, I determined that a case demonstrated improvement over time if there was an increase in level across the three phases or if the case demonstrated an accelerating trend in the feedback phases compared to baseline. When analyzing the data of the core components I concluded that 6 cases (Cases A, B, C, G, H, & I) demonstrated treatment integrity improvement over time and 3 cases (Cases D, E, & F) remained about the same over time across all three phases. Again, of the 6 cases demonstrating improvement, only 2 cases (Cases B & G) had declining baseline phases. No case showed a trend toward declining over time. When analyzing the data of the tips and reminders I concluded that 5 cases (Cases C, D, E, G, & I) demonstrated treatment integrity improvement over time, 3 cases (Cases B, F, & H) demonstrated a decline over time, and 1 case (Case A) remained about the same over time. Of the cases that demonstrated improvement, 3 cases had declining baseline trends (Cases E, G, & I).

**Parent Participant Feedback**

Overall, parental ratings of performance feedback, as assessed with the Intervention
Rating Profile-Performance Feedback (IRP-PF), reflected favorable opinions. The items on the IRP-PF and corresponding means and standard deviations are listed in Table 21.

Table 21

*Parent Participants Acceptability of Performance Feedback*

<table>
<thead>
<tr>
<th>Items (1=strongly disagree to 6=strongly agree)</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weekly verbal performance feedback is an acceptable way to help parents accurately use interventions (reading programs) with their children.</td>
<td>5.67</td>
<td>0.50</td>
</tr>
<tr>
<td>Weekly verbal performance feedback should prove effective in improving the accuracy with which parents implement intervention programs.</td>
<td>5.67</td>
<td>0.50</td>
</tr>
<tr>
<td>I would suggest the use of weekly verbal performance feedback for use with other parents.</td>
<td>5.78</td>
<td>0.44</td>
</tr>
<tr>
<td>My inaccurate implementation of the intervention was severe enough to warrant the use of weekly verbal performance feedback.</td>
<td>3.22</td>
<td>1.39</td>
</tr>
<tr>
<td>Most parents would find weekly verbal performance feedback suitable for addressing inaccurate program implementation.</td>
<td>5.13</td>
<td>0.64</td>
</tr>
<tr>
<td>I would be willing to receive verbal performance feedback again in the future.</td>
<td>5.67</td>
<td>0.50</td>
</tr>
<tr>
<td>Verbal performance feedback would not result in negative side effects for the parent.</td>
<td>5.56</td>
<td>0.53</td>
</tr>
<tr>
<td>Verbal performance feedback is a fair way to handle my inaccurate intervention implementation.</td>
<td>5.56</td>
<td>0.53</td>
</tr>
<tr>
<td>I like the procedures (verbal performance feedback) used to assist me in implementing the intervention with my child.</td>
<td>5.56</td>
<td>0.53</td>
</tr>
<tr>
<td>Overall, the procedures used (verbal performance feedback) would be beneficial for parents.</td>
<td>5.56</td>
<td>0.53</td>
</tr>
</tbody>
</table>
When asked to rate self-evaluation as a feedback method, assessed via the Intervention Rating Profile-Self Evaluation (IRP-SE), parents rated this mode of feedback favorably. The items on the IRP-SE and corresponding means and standard deviations are listed in Table 22.

Table 22

*Parent Participants Acceptability of Self-Evaluative Feedback*

<table>
<thead>
<tr>
<th>Items (1=strongly disagree to 6=strongly agree)</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weekly self-evaluation feedback is an acceptable way to help parents accurately use interventions (reading programs) with their children.</td>
<td>5.22</td>
<td>0.44</td>
</tr>
<tr>
<td>Weekly self-evaluation feedback should prove effective in improving the accuracy with which parents implement intervention programs.</td>
<td>5.22</td>
<td>0.44</td>
</tr>
<tr>
<td>I would suggest the use of weekly self-evaluation feedback for use with other parents.</td>
<td>5.22</td>
<td>0.44</td>
</tr>
<tr>
<td>My inaccurate implementation of the intervention was severe enough to warrant the use of weekly self-evaluation feedback.</td>
<td>3.11</td>
<td>1.54</td>
</tr>
<tr>
<td>Most parents would find weekly self-evaluation feedback suitable for addressing inaccurate program implementation.</td>
<td>4.78</td>
<td>0.67</td>
</tr>
<tr>
<td>I would be willing to receive self-evaluation feedback again in the future.</td>
<td>5.33</td>
<td>0.71</td>
</tr>
<tr>
<td>Self-evaluation feedback would not result in negative side effects for the parent.</td>
<td>5.44</td>
<td>0.73</td>
</tr>
<tr>
<td>Self-evaluation feedback is a fair way to handle my inaccurate intervention implementation.</td>
<td>5.33</td>
<td>0.87</td>
</tr>
<tr>
<td>I like the procedures (self-evaluation feedback) used to assist me in implementing the intervention with my child.</td>
<td>5.22</td>
<td>0.67</td>
</tr>
</tbody>
</table>
Overall, the procedures used (self-evaluation feedback) would be beneficial for parents.

Overall, parents reported high acceptability with using the HELPS Program. The items on the Intervention Rating Profile-Modified and corresponding means and standard deviations are listed in Table 23.

Table 23

*Parent Participants Acceptability of HELPS Program*

<table>
<thead>
<tr>
<th>Items (1=strongly disagree to 6=strongly agree)</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>This would be an acceptable intervention for a child’s reading problem.</td>
<td>5.50</td>
<td>0.53</td>
</tr>
<tr>
<td>This intervention should prove effective in changing a child’s reading skills.</td>
<td>5.37</td>
<td>0.52</td>
</tr>
<tr>
<td>I would suggest this intervention to other parents.</td>
<td>5.67</td>
<td>0.50</td>
</tr>
<tr>
<td>The child’s reading problem is severe enough to warrant the use of this intervention.</td>
<td>5.00</td>
<td>0.71</td>
</tr>
<tr>
<td>Most parents would find this intervention suitable for reading problems.</td>
<td>5.22</td>
<td>0.44</td>
</tr>
<tr>
<td>I would be willing to use this intervention at home.</td>
<td>5.67</td>
<td>0.50</td>
</tr>
<tr>
<td>This intervention would <em>not</em> result in negative side-effects for the child.</td>
<td>5.55</td>
<td>0.53</td>
</tr>
<tr>
<td>This intervention would be appropriate for a variety of children.</td>
<td>5.44</td>
<td>0.53</td>
</tr>
</tbody>
</table>
Table 23 (con’t)

<table>
<thead>
<tr>
<th>Description</th>
<th>Rating</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>This intervention is reasonable for the reading problem described.</td>
<td>5.22</td>
<td>0.44</td>
</tr>
<tr>
<td>I liked the procedures used in this intervention.</td>
<td>5.22</td>
<td>0.67</td>
</tr>
<tr>
<td>This intervention is a good way to handle the child’s reading problem.</td>
<td>5.11</td>
<td>0.60</td>
</tr>
<tr>
<td>Overall, this intervention would be beneficial for a child.</td>
<td>5.78</td>
<td>0.44</td>
</tr>
</tbody>
</table>

**Parent interview feedback.** Parents were interviewed at the completion of the study as a means of collecting qualitative feedback about the two forms of feedback and their experience using the HELPS reading program. All parents reported positive experiences using the HELPS Program and receiving feedback on their implementation of the program. All parents stated that the weekly observation and feedback session (regardless of feedback type) was helpful. All parents also said the performance feedback was helpful to improving implementation of the HELPS Program, and most parents said the Self-evaluation feedback was helpful to improving implementation of the HELPS Program, as two parents reported that they were less sure about its effectiveness given that they had already received performance feedback in their first experimental feedback phase and completed self-evaluation second. When given a choice, however, seven of the nine parents said they would prefer the performance feedback over self-evaluation.

Those parents who preferred performance feedback stated that they liked the explanations of the steps they missed and examples of how to do steps correctly. Parents liked receiving positive feedback in addition to the feedback on what was implemented.
incorrectly. Of the two parents who preferred the self-evaluation, one parent said it provided an objective “balance system.” All parents stated that the amount of time required to get feedback during either feedback phase was sufficient. Most parents did not feel that the weekly observation and feedback session was burdensome, and those who indicated that the weekly meetings were somewhat burdensome found it difficult primarily due to scheduling difficulties around work and other activities. All parents felt that the weekly observation and feedback session held them accountable to implementing the program with the prescribed frequency of three sessions per week.

The feedback regarding parents’ use of the HELPS Program was positive, as all but one parent felt the HELPS Program definitely helped their child improve their reading fluency. The one parent who expressed some hesitancy about the program’s effectiveness stated that her child’s performance fluctuated on some days. In general both the parents and their children enjoyed using the HELPS Program. Parents reported enjoying the one-on-one time with their child, liked the structure of the program, and liked that they had to provide specific praise and feedback to their child regarding their child’s reading behaviors. As per the parent feedback, student participants enjoyed the reward component of the HELPS Program. When asked generally what worked with the HELPS Program, parents felt the structured nature of the program and the repetition of the steps supported their child’s reading fluency development. There was varied feedback on the components of the program that were difficult to implement. This feedback including general difficulties, such as following the flowchart to ensure that all steps were implemented, and were implemented in the correct order, as well as specific difficulties, such as adhering to the scripted directions, correctly
marking incorrect words during timed readings, and remembering to place a bracket after the last work read in a minute during the timed readings. In general, however, parents found the program to be easy to use and reported greater ease of use over time.

**Student Outcomes**

I compared each student’s pre- and post-test reading achievement scores on the TOWRE-2, Sight Word Efficiency (SWE), TOWRE-2, Phonemic Decoding Efficiency (PDE), GORT-Fluency, and GORT-Comprehension by calculating a reliable change index (RC) (Jacobson & Truax, 1991) for each achievement measure. A significant RC index is a score of 1.96 or greater. The results are presented in Table 24.

<table>
<thead>
<tr>
<th>Case</th>
<th>TOWRE-2 SWE</th>
<th>TOWRE-2 PDE</th>
<th>GORT-Fluency</th>
<th>GORT-Comp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case A</td>
<td>-0.99</td>
<td>0</td>
<td>0.71</td>
<td>0</td>
</tr>
<tr>
<td>Case B</td>
<td>-0.56</td>
<td>1.24</td>
<td>0</td>
<td>-0.71</td>
</tr>
<tr>
<td>Case C</td>
<td>0</td>
<td>-1.24</td>
<td>0</td>
<td>-1.42</td>
</tr>
<tr>
<td>Case D</td>
<td>-0.14</td>
<td>-0.35</td>
<td>0</td>
<td>-0.71</td>
</tr>
<tr>
<td>Case E</td>
<td>1.13</td>
<td>0</td>
<td>0</td>
<td>-0.71</td>
</tr>
<tr>
<td>Case F</td>
<td>-1.55</td>
<td>-1.06</td>
<td>0.71</td>
<td>0</td>
</tr>
<tr>
<td>Case G</td>
<td>0</td>
<td>-0.18</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 24 (con’t)

<table>
<thead>
<tr>
<th>Case</th>
<th>0.71</th>
<th>-1.94</th>
<th>0.71</th>
<th>-1.42</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case</td>
<td>-1.13</td>
<td>2.65*</td>
<td>0.71</td>
<td>-0.71</td>
</tr>
</tbody>
</table>

Note. *significant at the p<.05 level.

In addition to analyzing student participants’ pre- and post-scores on measures of reading achievement, I also compared students’ fall benchmark score on the DIBELS Oral Reading Fluency (DORF) measure with the winter benchmark DORF score. I examined fall and winter scores by comparing the difference between the scores with expected grade-level growth. I calculated the expected growth for each student by subtracting the fall WCPM benchmark goal from the winter WCPM benchmark goal. I then added the standard error of measure (SEM) of the DORF Triad, specific to each grade level, to the difference between fall and winter benchmark goals. Second grade students exceeded their expected growth if their winter benchmark score exceeded 28 WCPM ([Winter 72] – [Fall 52] = 20 + [SEM 8.00] = 28 WCPM). Third grade students exceeded their expected growth if their winter benchmark score exceeded 23 WCPM ([Winter 86] – [Fall 70] = 16 + [SEM 7] = 23 WCPM). Case G, a fourth grade student, would have exceeded the expected growth if the winter benchmark score exceeded 21 WCPM ([Winter 103] – [Fall 90] = 13 + [SEM 8.53] = 21 WCPM); however, this student does not have a winter benchmark score. Of the eight students who had fall and winter benchmark scores, two students exceeded expected growth. The results are presented in Table 25.
Table 25

*Students Participants’ Pre- and Post-HELPΣ DORF Scores*

<table>
<thead>
<tr>
<th>Student</th>
<th>Pre-Test DORF Score (Fall Benchmark)</th>
<th>Post-Test DORF Score (Winter Benchmark)</th>
<th>Difference Between Pre-Test and Post-Test Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>46</td>
<td>47</td>
<td>+1</td>
</tr>
<tr>
<td>B</td>
<td>29</td>
<td>43</td>
<td>+14</td>
</tr>
<tr>
<td>C</td>
<td>60*</td>
<td>85*</td>
<td>+25</td>
</tr>
<tr>
<td>D</td>
<td>77*</td>
<td>90*</td>
<td>+13</td>
</tr>
<tr>
<td>E</td>
<td>38</td>
<td>46</td>
<td>+8</td>
</tr>
<tr>
<td>F</td>
<td>46</td>
<td>66</td>
<td>+20</td>
</tr>
<tr>
<td>G</td>
<td>39</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>H</td>
<td>44</td>
<td>84*</td>
<td>+40**</td>
</tr>
<tr>
<td>I</td>
<td>78*</td>
<td>145*</td>
<td>+67**</td>
</tr>
</tbody>
</table>

*Note.* *= At or above benchmark. **= Exceeded expected growth.
Discussion

I conducted the present study with the primary purpose of evaluating the effects of two forms of feedback (i.e., performance feedback from an external party and self-evaluation) to determine if one form of feedback was superior to the other in improving treatment integrity of a reading fluency program. When the majority of the visual analysis features supported a difference between the two conditions on a case-level, treatment integrity was higher in the performance feedback condition. However, when examining the difference in treatment integrity of the core components between the performance feedback phase and the self-evaluation phase across cases, I could not conclude that either feedback method was superior to the other. The same was true for comparing the treatment integrity of the tips and reminders between the performance feedback phase and self-evaluation phase. The subset of cases that had core components data and tips and reminders data that were sufficiently stable to compare and showed a clear difference between the compared conditions (as measured by the five possible features) was insufficient to detect a significant difference between the two feedback methods, even if such a difference exists.

The inability to answer my research questions due to an insufficient number of cases that showed stable data and a clear difference between compared phases leaves two interpretations of the results. First, there may truly be a difference in the efficacy of these two feedback methods, but the treatment integrity data were insufficiently stable and consistent to detect this outcome. If the effect is small, the detection of a difference would require a larger sample size than my sample size of 9 cases. Second, there may not be a difference in the efficacy of performance feedback and self-evaluation. However, determining that both
feedback methods are equally efficacious is not possible, as it would require proving the null hypothesis (i.e., proving that both methods are equally efficacious). Because statistical tests presume the null hypothesis (i.e., that both feedback methods produce the same outcomes), the most conservative interpretation of my findings is that I do not have sufficient evidence to show the two feedback methods to be different.

The results of the visual analysis also suggest that, regardless of the order of feedback methods, there was a general trend of treatment integrity improvement over time for the both the core components and the tips and reminders. Of the 9 cases, 6 cases demonstrated improvement in treatment integrity of the core components from baseline through phase three. Five cases demonstrated improvement in treatment integrity of the tips and reminders from baseline through phase three; however, 3 cases also demonstrated a decline across time.

The results across cases suggest that the method of feedback may not be as important as the necessity for feedback in general, and that continued feedback over time helps improve integrity (i.e., feedback has a cumulative effect on integrity). Although practice effects must be considered when analyzing the implementation of a program over time, as parents improved their treatment integrity in the absence of structured or prompted feedback, in all cases where parents demonstrated declining treatment integrity during the baseline phase, the introduction of a feedback method resulted in an immediate positive change in treatment integrity, and their treatment integrity generally improved over time.

In addition to comparing treatment integrity during the two feedback phases, I wanted to determine whether treatment integrity during each of the feedback phases was better than treatment integrity during the baseline phase. The results showed a statistically reliable
difference between the treatment integrity of the baseline data and the performance feedback data for the tips and reminders of the HELPS Program, supporting my hypothesis that treatment integrity would be higher in the performance feedback phase than in the baseline phase. The increase in treatment integrity following the introduction of performance feedback is consistent with previous findings from studies that examined the effects of performance feedback (Codding et al., 2008; Codding et al., 2005; Jones et al., 1997; Mortenson & Witt, 1998; Noell, et al., 1997; Witt et al., 1997).

Although I was able to detect a difference in the treatment integrity of the tips and reminders between the baseline and performance feedback data, I did not have a sufficient number of cases that showed a difference between the baseline and performance feedback treatment integrity of the core components. Thus, the inconclusive outcome to the comparison of the treatment integrity of the core components during baseline and performance feedback does not provide evidence for me to support my hypothesis.

Previous research has shown that the use of self-evaluation improved treatment integrity when compared to baseline (Belfiore et al., 2008; Pelletier et al., 2010; Sutherland & Wehby, 2001). I failed to find this outcome in my study, however, as I did not have a sufficient number of cases that demonstrated a difference between the baseline and self-evaluation, as determined by a majority of the visual analysis features favoring self-evaluation over baseline. Again, without a sample that has stable data and clear differences in the comparisons between phases to adequately compare self-evaluation with baseline treatment integrity, I do not have evidence to support my hypothesis.
Student Outcomes

Although the primary purpose of this study was to compare two forms of feedback, I also examined student outcomes. Overall, the comparison of students’ pre-HELPS scores with post-HELPS scores on measures of reading achievement did not reflect significant improvement in reading fluency or comprehension, as only 1 of 9 students demonstrated a statistically reliable change on one of the outcome measures (TOWRE-2-PDE). These findings are surprising given the research base supporting the efficacy of the HELPS Program. The comparison of students’ fall and winter DORF scores, however, reveal that all students improved their WCPM score from fall to winter. Of the 9 students, 4 students were at or above benchmark based on winter DORF scores, and 2 of the 4 students exceeded expected growth from fall to winter. Therefore, there is no consistent evidence to show that the HELPS Program boosted reading skills beyond what would be expected from traditional instruction in the time frame within my study.

Practical Implications

Given that I could not determine if performance feedback was superior to self-evaluation in improving treatment integrity, I cannot conclude that the components of the performance feedback condition made this feedback method any more effective than self-evaluation feedback. However, there are several variables pertinent to the use of performance feedback and self-evaluation that should invite future research on the practical implications of these two feedback methods, including personnel resources, potential observer effects, and parent preference for feedback method.

Self-evaluation and performance feedback have differing demands for personnel
resources. On the one hand, self-evaluation requires less personnel resources, as the parent needs only a tape player and an implementation protocol form to complete. No third-party observer is required to obtain feedback. Eliminating a direct observer from an intervention session may also be considered less invasive and may help preserve parents’ feelings of competency. On the other hand, performance feedback requires an observer to be present during some of the sessions, which requires scheduling and possible travel if observing the parent in the home, in addition to the time directly observing and providing feedback. Thus, performance feedback is more time-intensive than self-evaluation, and therefore requires more personnel resources.

A possible limitation to self-evaluation, however, is ensuring feedback completion. Without an observer present during the self-evaluation process, it is unclear whether the parent would complete the task. Also, self-evaluation requires more time on the part of the parent, as the parent must conduct the session with the child and then listen to the entire session again while completing the implementation protocol. When parents completed their self-evaluations, I was present as they listened to the tape and completed the intervention protocol form. It is possible that my presence during and after the session may have led parents to more regularly and comprehensively complete the self-evaluation process. Although parents in this study did not report that the additional time following the session to complete the self-evaluation process was burdensome, they may have been more motivated to complete the feedback process than the typical parent, given that they volunteered to participate in this study. It is also possible that parents were less inclined to share with me that they did not want to complete the self-evaluation process and only did so because I was
present during the session observing the parent. Observer effects on parents’ self-evaluation behavior could be explored in additional research; my design did not allow me to compare self-evaluation with and without an observer (i.e., me) being present.

Parents reported positive responses to performance feedback and to self-evaluation. Positive ratings of performance feedback and self-evaluation are consistent with previous research findings (Codding et al., 2008; Codding et al., 2005; Noell et al., 2005; Noell et al., 2002; Sutherland & Wehby, 2001); however, previous studies assessed educators’ experiences with performance feedback, not parents’ experiences. There is limited research that examines the parents’ role as interventionist receiving performance feedback, and those that used parents as interventionists did not include measures of feedback acceptability. Thus, the findings from this study provide preliminary support for feedback acceptability, regardless of whether it was supplied by an observer or through self-evaluation. Parents also reported that both feedback methods assisted with improving treatment integrity. When asked what method was preferred, however, the majority of parents stated that they preferred performance feedback over self-evaluation. This is important to note, as this form of feedback required an observer to be present during the session and required additional time following the session to provide the feedback.

In summary, a professional should consider several factors when deciding whether to use self-evaluation or performance feedback to help parents achieve high rates of treatment integrity. When considering the use of self-evaluation, the professional who is working with the parent must ensure that the parent is completing the self-evaluation process accurately and adhering to a schedule for completing the self-evaluation process. In this study, parents
tended to inflate their self-evaluation treatment integrity scores by endorsing the implementation of components of the program that they did not accurately implement, or implement at all. Thus, self-rating scores were not accurate (i.e., they were inflated). Thus, although the self-evaluation process allows the parent to be more independent in gathering feedback, particularly as a check on the presence/absence of completed steps, it may be necessary to meet with the parent to clarify intervention components that the parent endorses as completing correctly, as the parent may not be able to self-correct inaccurately completed components.

When considering the use of performance feedback, observers must schedule appointments with the interventionist that provide sufficient time to observe an intervention session as well as provide the vocal feedback following the session. The feedback sessions must allow time to review correctly implemented intervention components, components that were incorrectly implemented or skipped, answer interventionist questions, review permanent products, and possibly demonstrate how to implement components of the intervention. Consequently, the professional must consider that using performance feedback will take more, and longer, sessions with parents than will self-evaluation.

Limitations

Several limitations were present in this study. First, I was unable to answer my research questions and three of my hypotheses due to a combination of data limitations. These data limitations included several cases had data that were insufficiently stable to compare, and were thus, discarded from the analyses, and several cases had data that were sufficiently stable to compare but resulted in ties, that is, no clear difference between the
compared phases. The combination of data instability and ties prevented me from detecting a statistically significant difference between the integrity during the performance feedback phase and the integrity during the self-evaluation phase. Therefore, it is not possible for me to determine whether performance feedback provided by a third party is better or worse than self-evaluation in influencing intervention integrity in parents.

Second, as is true of single-case designs, the number of data points per phase can influence the reliability of the data level and trend. Although all phases across all cases had at least three data points per phase, some cases had fewer data points within a phase, and those data fluctuated greatly. Although lengthening the phase would not have guaranteed a stable level and/or trend, longer intervals should provide a more reliable estimate of level and/or trend, which would strengthen the validity of the data analysis.

Third, several cases had an accelerating trend line in the baseline phase (5 cases for analysis of core components; 3 cases for the tips and reminders). When a case had an accelerating baseline phase it was difficult to determine if improved treatment integrity following the introduction of a feedback method was related to the feedback method. Some parents demonstrated an immediate increased treatment integrity level change with the introduction of feedback, whereas others decreased their treatment integrity. Thus, the results for introducing feedback following an accelerating baseline trend were not consistent across cases. It is important to note, however, that treatment integrity immediately improved with the introduction of feedback following a decelerating baseline trend.

Fourth, I compared data in the last phase of each case with baseline data, which is not typically done in single case design studies without another baseline phase between treatment
phases. I did not include a second baseline phase because I was concerned attrition would increase if I lengthened the study. Plus, the benefit of administering the two feedback methods in successive phases allowed me to compare possible changes in treatment integrity when moving from the first feedback method to the second feedback method, and comparing the two feedback methods was the primary purpose of the study. Given that I do not have a second baseline in which to compare the third phase, however, I must use caution when interpreting my results of comparing baseline with the third feedback phase.

Fourth, the accuracy with which participants completed the implementation protocol during the self-evaluation phase was not particularly high. Parents tended to inflate their implementation score, endorsing their implementation of core components and tips and reminders that were not implemented during the session. It is possible that parents did not fully understand all of the core components and tips and reminders of the program and assumed they did each step and tip and reminder correctly, despite listening to the recorded session when completing the implementation protocol. Therefore, it is possible that the effect of self-evaluation was limited due to incomplete understanding of the program components and tips and reminders.

**Future Directions and Practice**

This study provides several future research directions. First, researchers should consider comparing performance feedback with self-evaluation to determine if these feedback methods differ in their efficacy on improving treatment integrity of the HELPS Program. If the difference between these methods is small, detecting this difference would require a study with a larger sample size than the study I conducted—but it also raises the
question whether it is worth the time and expense of doing research with larger samples simply to detect small (and perhaps clinically irrelevant) differences between methods.

Second, in this study parents completed the self-evaluation procedures only during the sessions that I directly observed. A logical extension of this study would be to evaluate observer effects on feedback completion by asking parents to complete self-evaluations with and without an observer present to determine if feedback completion differed with the presence/absence of a direct observer.

Third, the general feedback schedule for this study was such that I provided performance feedback or parents completed self-evaluation (depending on phase) immediately following one session per week, and the schedule for feedback remained at one feedback session per week for the entirety of the study. Regardless of parents’ integrity, I did not change the frequency of feedback meetings. In prior studies, researchers examined schedules for providing performance feedback in relation to treatment integrity level, with differing outcomes across studies (Codding et al., 2005; DiGennaro et al., 2007; DiGennaro et al., 2005; Mortenson & Witt, 1998; Noell et al., 2002; Noell et al., 2000; Witt et al., 1997). Given the lack of consensus regarding the optimal number of feedback sessions per week, or how the frequency of feedback sessions should change in response to changes in treatment integrity, additional research is warranted.

Performance feedback and self-evaluation are both feedback methods used by researchers and professionals to improve the treatment integrity of home-based academic program. Professionals should consider the merits and limitations of each feedback method to decide how to maximize effectiveness and efficiency.
Conclusion

Performance feedback has been shown through research to improve treatment integrity. Typically, this performance feedback has been delivered in the form of vocally-mediated feedback delivered by a third party (e.g., consultant, coach, teacher). Performance feedback (i.e., third party vocal feedback) vs. self-evaluation feedback conditions have not been previously compared to determine if one form of feedback method is superior to the other for improving treatment integrity. This is the first study to date that has directly compared performance feedback to self-evaluation. However, data instability, coupled with ties, failed to produce enough cases to adequately test my research questions, and three of my four hypotheses. Although my comparison of performance feedback to self-evaluation is inconclusive, the study provides future directions for researchers seeking to better understand how differing feedback methods may improve implementation integrity.

Comparing both feedback methods with baseline yielded one reliable difference: treatment integrity of the tips and reminders was better in the performance feedback condition than in the baseline condition. Additional comparisons between baseline and the two feedback methods could not be conducted due to data instability and a high number of ties. However, in cases where parents had declining baseline treatment integrity, treatment integrity improved following the introduction of feedback. Additionally, parents demonstrated a general trend of treatment integrity improvement over time regardless of the order of feedback, suggesting that providing feedback over time may be more important for improving integrity than the type of feedback provided. Furthermore, parents rated both
performance feedback and self-evaluation favorably. If given a preference, however, most parents preferred performance feedback.
REFERENCES


Good, R.H., & Kaminski, R.A., (with Cummings, K., Dufour-Martel, C., Petersen, K.,...


Schulte, A.C., Easton, J.E., & Parker, J. (2009). Advances in treatment integrity research:


Appendix A

HELPS Parent Project
CONFIDENTIAL SCREENING QUESTIONNAIRE

Information requested on this questionnaire will be helpful in describing your child for our reading project. Please feel free to add as much information as you want. You may find that some questions do not apply to your child. If this is the case, please write N/A (Not Applicable) in that space. The highest standards of professional confidentiality are maintained. Information about any particular child can be released only with the explicit written consent of their parent or legal guardian.

Name: _________________________________ DOB: ________________

Child’s Name: ________________________ Child’s DOB: ____________

Relationship to Child: __________________

I. BASIC DEMOGRAPHIC INFORMATION OF CHILD

Age of Child: _____ Sex: _____ Grade: _____ School_____________________

Hand Child Writes With: _______________ Ethnicity: ______________________

First language child learned to speak: ___________ Second Language ___________

Primary language spoken in your child’s home: ____________________________

II. MEDICAL/EDUCATIONAL HISTORY OF CHILD

Please indicate if this child has ever had any of the following conditions by answering YES or NO. If YES, please specify when and for how long:

_____ 1. Problems with speech (e.g., stuttering) or hearing?

   Specify:
   ________________________________________________________________

   Has your child’s speech or hearing problem been corrected?
   ________________________________________________________________

_____ 2. Problems with vision?

   Specify:
Has your child’s vision problem been corrected?

3. Has this child received extra support at school for reading? If yes:
   Grade(s): ______________

4. Has this child ever been retained in school (i.e., repeated a grade)? If yes:
   Grade repeated: ______________

III. INFORMATION ABOUT PARENTS/GUARDIANS:
1. Please circle the term that represents your relationship to the child:
   - Biological Mother
   - Adoptive Mother
   - Stepmother
   - Grandmother
   - Biological Father
   - Adoptive Father
   - Stepfather
   - Grandfather
   - Other Guardian (Please specify): ______________

2. What is your marital status? ______________

3. If married, is your spouse the child’s biological mother/father? ______________

Please fill out the chart regarding information about yourself (and, if applicable, spouse):

<table>
<thead>
<tr>
<th>Years of Education</th>
<th>Mother/Legal Guardian</th>
<th>Father/Legal Guardian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highest Level of Formal Schooling: (e.g., Some High School, High School (GED), Vocational Certificate/Degree, BA/BS, MS/MA, Ph.D.)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Occupation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Job Title</td>
</tr>
</tbody>
</table>
IV. INFORMATION ABOUT READING ACTIVITIES

Please use the following scale to answer questions about home activities and your child. Please read each question carefully and circle the answer that is most reflective.

4 = Almost daily  3 = Once or twice a week  2 = Once or twice a month  1 = Hardly ever

1) How often does your child see an adult read a book/newspaper/magazine?
   4   3   2   1

2) How often does your child see an adult read on a computer?
   4   3   2   1

3) How often does your child ask for help with homework?
   4   3   2   1

4) How often do you monitor your child’s homework?
   4   3   2   1

5) How often do you read aloud to your child?
   4   3   2   1

6) How often does your child read aloud a book to you?
   4   3   2   1

Listed below are activities your child may do. It is not expected that children do all these activities, but we are interested in how often your child has participated in each activity in the last year. Please circle the rating that best applies to your child using the following scale:

4 = Almost daily  3 = Once or twice a week  2 = Once or twice a month  1 = Hardly ever

1) Visits the community library and brings books home.
   4   3   2   1

2) Reads signs, labels, grocery lists.
   4   3   2   1

3) Plays educational games (e.g. Scrabble).
   4   3   2   1

4) Plays computer games involving reading.
   4   3   2   1

5) Writes a story/note/card/diary.
   4   3   2   1

6) Reads recipe/craft/game instructions with you.
   4   3   2   1

7) Reads a book for fun (i.e. was not asked to read the book by an adult).
   4   3   2   1

8) Watches educational TV shows with an adult.
   4   3   2   1
Parents sometimes help their children while reading books together. Using the same scale, please indicate if you use any of these reading activities with your child.

4 = Almost daily  
3 = Once or twice a week  
2 = Once or twice a month  
1 = Hardly ever

1) Teach words on flashcards.  
   4 3 2 1

2) Ask your child to read aloud for practice.  
   4 3 2 1

3) Assist with new word by sounding out letters.  
   4 3 2 1

4) Assist with new word by using the meaning of the sentence.  
   4 3 2 1

5) Assist with new word by asking child to try the word again.  
   4 3 2 1

6) Assist with new word by providing the correct word.  
   4 3 2 1

Please circle the answer that best describes your home:

1) About how many hours of television/ videos does your child watch each day?  
   > 5 hours  3-5 hours  1-2 hours  < 1 hour  0 hours

2) About how many books would you estimate that you have in your home?  
   > 500  300-500  100-300  50-100  < 50
Appendix B

HELPs One-on-One Program: Implementation Protocol

Steps below that are denoted with an asterisk (*) should be implemented in combination with the “HELPs One-on-One Program: Scripted Directions.”

1. *Parent reads introductory statements and expectations – includes Verbal Cuing procedure (15 seconds)
   - Overall: goal of program is for student to do his/her best reading. This means he/she tries to read quickly, accurately, and with good expression. Also, the student tries to remember what happens in the story and tries to remember the difficult words that he/she practices.

2. *Student Timed Reading with Passage A, as indicated on his/her Progress Tracking Form (1 to 1.5 minutes)
   - NOTE: Students who meet the WCPM criterion can be stopped at 1 minute. Students who do not meet the WCPM criterion should read 1.5 minutes to allow for some additional practice (though as indicated in the scripted directions, the student’s last word read at one minute should be indicated with a bracket).

3. *Parent asks student to say what he/she remembers about the story (Retell Check), asking the student to recall what happened in appropriate sequential order (30 to 45 seconds)

**IF STUDENT MEETS THE READING GOAL** (See back page if student does not meet the Reading Goal)

The Goal is met when the student (a) meets the WCPM criterion, (b) meets the WIPM criterion, and (c) can adequately recall parts of the story. (See Table on back of sheet for all Goal criteria according to student grade level).

4a. When the student meets the Reading Goal, the parent should: (15 seconds)
   - Provide praise for meeting the Goal and immediately graph the student’s performance on Passage A.
   - Tell the student he/she will earn at least one star at the end of the session for meeting the Goal.
   - Obtain the next story in the HELPs Curriculum.

5a. *Student Timed Reading of next story (Passage B) in the HELPs Curriculum (1 minute)

6a. *Parent implements phrase-drill error correction on all incorrectly read words (15 to 45 seconds)

7a. *Student Timed Reading of Passage B a second time for no more than one minute (1 minute)

8a. *Parent models fluent oral reading of Passage B while student follows along (1 to 1.5 minutes)
9a. *Student Timed Reading of Passage B a third time for no more than one minute (1 minute)

10a. Parent graphs WCPM and WIPM for the student’s first and third reading of Passage B (15 seconds)
    □ While doing so, parent provides specific, ENTHUSIASTIC praise (and feedback) regarding the student’s reading and praises student for reading improvements, if applicable.

11a. Parent awards stars on the student’s Star Chart (15 to 30 seconds)
    □ Parent gives one star because student met the Reading Goal with Passage A.
    □ Parent gives a second star if student clearly demonstrates effort when practicing Passage B and reads more WCPM during the last reading compared to the first reading of Passage B.

12a. Parent records all information on the student’s Progress Tracking Form and indicates which passage the student should read at the start of the next session (30 seconds)

(Estimated time of implementation if student does meet Reading Goal: 7.5 to 9 minutes)

IF STUDENT DOES NOT MEET THE READING GOAL

4b. *Parent models fluent oral reading of Passage A while student follows along (1 to 1.5 minutes)

5b. *Student Timed Reading of Passage A a second time for no more than one minute (1 minute)

6b. *Parent implements phrase-drill error correction on all incorrectly read words (15 to 45 seconds)

7b. *Student Timed Reading of Passage A a third time for no more than one minute (1 minute)

8b. *Parent implements phrase-drill error correction on all incorrectly read words (15 to 45 seconds)

9b. Parent graphs WCPM and WIPM for the student’s first and third reading of Passage A (15 seconds)
    □ While doing so, parent provides specific, ENTHUSIASTIC praise (and feedback) regarding the student’s reading and praises student for reading improvements, if applicable.

10b. Parent awards stars on the student’s Star Chart (15 to 30 seconds)
    □ Parent gives one star if student clearly demonstrates effort when practicing Passage A and reads more WCPM during the last reading compared to the first reading.

11b. Parent records all information on the student’s Progress Tracking Form and indicates which passage the student should read at the start of the next session (30 seconds)
(Estimated time of implementation if student does not meet Reading Goal: 7 to 9 minutes)

Reading Goals According to the Student’s Grade Level

<table>
<thead>
<tr>
<th>Grade Level</th>
<th>WCPM with Passage A</th>
<th>WIPM with Passage A</th>
<th>Retell Check with Passage A</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Grade</td>
<td>80 or more</td>
<td>3 or less</td>
<td>Adequately retells story*</td>
</tr>
<tr>
<td>Second Grade</td>
<td>100 or more</td>
<td>3 or less</td>
<td>Adequately retells story*</td>
</tr>
<tr>
<td>Third Grade</td>
<td>120 or more</td>
<td>3 or less</td>
<td>Adequately retells story*</td>
</tr>
<tr>
<td>Fourth Grade</td>
<td>135 or more</td>
<td>3 or less</td>
<td>Adequately retells story*</td>
</tr>
</tbody>
</table>

* For example, student retells parts of the story for at least 30 seconds or otherwise correctly states names of characters and major events in the story. Retell of the story in the correct sequential order of major events is encouraged but not required to pass the Retell Check.
Observation Checklist for Implementing the HELPS One-on-One Program

Parent observed: ______________________________  Student: ______________________________
Observer: ____________________________________  Date: _______________________________

Implementation of Core Procedures

(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)

List of Steps (if Goal is met): 1 ☑; 2 ☑; 3 ☑; 4a ☑; 5a ☑; 6a ☑; 7a ☑; 8a ☑; 9a ☑; 10a ☑; 11a ☑; 12a ☑
List of Steps (if Goal is not met): 1 ☑; 2 ☑; 3 ☑; 4b ☑; 5b ☑; 6b ☑; 7b ☑; 8b ☑; 9b ☑; 10b ☑; 11b ☑

Percentage of steps completed = total steps completed / total steps possible _____/_____ x 100: ______%

Implementation of Tips and Reminders (steps missed below are considered minor errors)

(Place “✓” in the box for completed steps; Place “X” in the box for skipped steps; Place circle “O” around the box for non-applicable [NA] steps)

General Implementation Procedure
☐ Parent 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, Star Chart, Bonus Bag, Implementation Flow Chart, and Scripted Directions. Also, the prize box was reasonably accessible.
☐ Parent used Scripted Directions or Abbreviated Directions as advised at top of Abbreviated Directions.

Repeated Reading (and Timed Reading) Procedure
☐ After each student reading, parent indicated on examiner passage (w/a bracket) the # of words read in one min
If parent recorded all student readings with dry-erase marker before transferring scores to Progress Tracking Form, he:
☐ Put the appropriate number (i.e., 1, 2, or 3) next to the one-minute bracket.
☐ Marked student errors differently during each reading (e.g., first reading = slash, second reading = underline, third reading = circle).

Retell Check Procedure
☐ Before prompting student to begin the Retell Check, parent made sure student could not review the passage during the Retell Check.
☐ Parent used broad follow-up questions to solicit student’s retell only if student was unable to retell the passage for approximately 30 seconds.
☐ Parent implemented Retell Check for no more than 45 seconds unless he/she made a decision prior to the session to lengthen the Retell Check.

Goal Setting Procedure
☐ Parent told student he/she met the Reading Goal.

Phrase-Drill Error Correction Procedure
☐ Parent asked student to practice “logical” phrases.
☐ Parent told student to “READ” the phrases, and did not ask the student to “SAY” or “REPEAT” phrases.
☐ Parent had student practice all incorrectly read words (up to 5 or until time permits)
☐ Parent pointed (or had the student point) to each word practiced.
If student practiced words that were read correctly but less fluently (see above step), parent explained to student that he/she read the words correctly, but will practice them because they are difficult.

If student made 1 or fewer errors, the parent told the student to practice 1-3 words or phrases that were read less fluently.

**Modeling Procedure**
- Parent read aloud at a pace just a little faster than the student’s reading ability.
- Parent read with good expression.
- Parent read at a volume the student could clearly hear.
- Parent paused 5-7 times to have student read the next word in the passage.

**Performance Feedback (Graphing) Procedure**
- While graphing, parent gave verbal feedback and praise regarding the student’s WCPM and WIPM scores.
- Parent graphed WCPM and WIPM on 2 or 3 readings (3 if the Goal was met; 2 if the Goal was not met).
- Parent connected lines between WCPM (and WIPM) scores only for scores of the same passage.
- Parent circled the data point and session number when the student began a new passage.

**Motivational (Reward) Procedure**
- Throughout the session, parent provided a minimum of three different praise statements regarding student’s reading behavior.
- When awarding stars on Star Chart, parent accurately told student why he/she earned each star.
- With enthusiasm, parent 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 student for specific reading behaviors or improvements at the end of the session.
- If the student landed on OR passed a shaded square on Star Chart, student was allowed to select a ticket from the bonus bag and parent correctly recorded the bonus stars written on the ticket.
- Parent conveyed that improved reading skills, rather than the opportunity to earn stars/prizes, is the primary reason the student should put forth effort during each HELPS session.

**Using the Progress Tracking Form**
- After finishing session, parent completed Progress Tracking Form before erasing data from examiner passage.
- Parent recorded 2 or 3 sets of WCPM/WIPM scores on the Progress Tracking Form, as determined by whether the student met his/her Reading Goal on passage A (3 sets of scores were recorded if Goal was met; 2 sets of scores were recorded if Goal was not met).
- Parent recorded relevant information in the Notes column of the Progress Tracking Form (e.g., student difficulties with Retell Check, behavior problems, attention difficulties, etc.).

Total steps applicable = 31 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 the Student’s Timed Readings**

Discrepancies / Total words read:

Reading 1: ____/ ____ ISRA%:______
Reading 2: ____/ ____ ISRA%:______

Reading 3: ____/ ____ ISRA%:______
Reading 4: ____/ ____ ISRA%:______
(When applicable)
Appendix D

Observation Summary Form (OSF)

Parent: ____________________________  Child: __________________________

RA: ____________________________  Date: __________________________


Student’s session #: _______  Student met goal (circle):  Yes  No

Duration of intervention session: _____ minutes _____ seconds

Parent’s enthusiasm with student: _______  Organization: _______
(Rate Enthusiasm and Organization/preparation 1 – 5; 1=poor, 3=average, 5=outstanding)

Parent implemented _______ % of primary protocol and _______ % of tips/reminders.

Core procedures not implemented (List step numbers and write notes if needed):
______________________________________________________________________________
______________________________________________________________________________
______________________________________________________________________________
______________________________________________________________________________

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

Questions or concerns raised by parent (or additional notes or feedback provided by observer):
______________________________________________________________________________
______________________________________________________________________________
______________________________________________________________________________

Total duration of feedback session: _____ minutes _____ seconds

**Only for verbal feedback phase:
Were all parent questions/concerns addressed:  Yes  No  N/A

Were all missed steps & tips reviewed:  Yes  No  N/A
## Appendix E

**Average HELPS session length, length of feedback sessions, & enthusiasm and organization ratings**

<table>
<thead>
<tr>
<th>Participant</th>
<th>Length of HELPS Session</th>
<th>Length of Feedback Session</th>
<th>Enthusiasm</th>
<th>Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline*</td>
<td>Self-Eval</td>
<td>PF</td>
<td>Scale: 1 – 5</td>
</tr>
<tr>
<td>A</td>
<td>21:09</td>
<td>05:00</td>
<td>21:24</td>
<td>18:08</td>
</tr>
<tr>
<td>B</td>
<td>18:02</td>
<td>04:53</td>
<td>21:38</td>
<td>21:00</td>
</tr>
<tr>
<td>C</td>
<td>15:16</td>
<td>01:00</td>
<td>21:00</td>
<td>08:12</td>
</tr>
<tr>
<td>D</td>
<td>14:56</td>
<td>03:04</td>
<td>19:59</td>
<td>08:41</td>
</tr>
<tr>
<td>E</td>
<td>16:50</td>
<td>01:00</td>
<td>20:38</td>
<td>10:23</td>
</tr>
<tr>
<td>F</td>
<td>18:05</td>
<td>01:00</td>
<td>22:42</td>
<td>06:17</td>
</tr>
<tr>
<td>G</td>
<td>17:11</td>
<td>01:48</td>
<td>22:49</td>
<td>11:22</td>
</tr>
<tr>
<td>H</td>
<td>11:52</td>
<td>03:10</td>
<td>14:52</td>
<td>05:11</td>
</tr>
<tr>
<td>I</td>
<td>14:57</td>
<td>02:29</td>
<td>17:18</td>
<td>11:23</td>
</tr>
</tbody>
</table>

*Note.* Although feedback was not provided during this phase, parents had the opportunity to ask questions, which were then recorded on the Observation Summary Form. All times reported in minutes:seconds. Scale: 1 = Poor, 3 = Average, 5 = Outstanding.
Appendix F

Intervention Rating Profile – (IRP-15)- Modified

Name: _______________________________   Date: _____________________

Please rate the intervention (HELPS Program) along the following dimensions. Please circle the number which best describes your agreement or disagreement with each statement.

1= strongly disagree   2= disagree   3= slightly disagree   4= slightly disagree   5= agree   6= strongly agree

1. This would be an acceptable intervention for a child’s reading problem.  1  2  3  4  5  6

2. This intervention should prove effective in changing a child’s reading skills.  1  2  3  4  5  6

3. I would suggest this intervention to other parents.  1  2  3  4  5  6

4. The child’s reading problem is severe enough to warrant use of this intervention.  1  2  3  4  5  6

5. Most parents would find this intervention suitable for reading problems.  1  2  3  4  5  6

6. I would be willing to use this intervention at home.  1  2  3  4  5  6

7. This intervention would not result in negative side-effects for the child.  1  2  3  4  5  6

8. This intervention would be appropriate for a variety of children.  1  2  3  4  5  6

9. This intervention is reasonable for the reading problem described.  1  2  3  4  5  6

10. I liked the procedures used in this intervention.  1  2  3  4  5  6

11. This intervention is a good way to handle this child’s reading problem.  1  2  3  4  5  6

12. Overall, this intervention would be beneficial for a child.  1  2  3  4  5  6
Appendix G

Intervention Rating Profile – Verbal Feedback

Name: _______________________________  Date: __________________________

Received verbal performance feedback 1st: YES or NO (circle)

The purpose of this questionnaire is to get information that will help school psychologists in their work with parents. Please circle the number which best describes your agreement or disagreement with each statement.

1- strongly disagree  2-disagree  3-slightly disagree  4-slightly agree  5-agree  6-strongly agree

Weekly verbal performance feedback is an acceptable way to help parents accurately use interventions (reading programs) with their children.  1  2  3  4  5  6

Weekly verbal performance feedback should prove effective in improving the accuracy with which parents implement intervention programs.  1  2  3  4  5  6

I would suggest the use of weekly verbal performance feedback for use with other parents.  1  2  3  4  5  6

My inaccurate implementation of the intervention was severe enough to warrant the use of weekly verbal performance feedback.  1  2  3  4  5  6

Most parents would find weekly verbal performance feedback suitable for addressing inaccurate program implementation.  1  2  3  4  5  6

I would be willing to receive verbal performance feedback again in the future.  1  2  3  4  5  6

Verbal performance feedback would not result in negative side effects for the parent.  1  2  3  4  5  6

Verbal performance feedback is a fair way to handle my inaccurate intervention implementation.  1  2  3  4  5  6

I like the procedures (verbal performance feedback) used to assist me in implementing the intervention with my child.  1  2  3  4  5  6

Overall, the procedures used (verbal performance feedback) would be beneficial for parents.  1  2  3  4  5  6

Comments regarding verbal performance feedback:

_________________________________________________________________________________________
Appendix H

Intervention Rating Profile – Self-Evaluation

Name: _______________________________  Date: _______________________

Received self-evaluation feedback 1st: YES or NO (circle)

The purpose of this questionnaire is to get information that will help school psychologists in their work with parents. Please circle the number which best describes your agreement or disagreement with each statement.

1- strongly disagree  2-disagree  3-slightly disagree  4-slightly agree  5-agree  6-strongly agree

Weekly self-evaluation feedback is an acceptable way to help parents accurately use interventions (reading programs) with their children.

Weekly self-evaluation feedback should prove effective in improving the accuracy with which parents implement intervention programs.

I would suggest the use of weekly self-evaluation feedback for use with other parents.

My inaccurate implementation of the intervention was severe enough to warrant the use of weekly self-evaluation feedback.

Most parents would find weekly self-evaluation feedback suitable for addressing inaccurate program implementation.

I would be willing to receive self-evaluation feedback again in the future.

Self-evaluation feedback would not result in negative side effects for the parent.

Self-evaluation feedback is a fair way to handle my inaccurate intervention implementation.

I like the procedures (self-evaluation feedback) used to assist me in implementing the intervention with my child.

Overall, the procedures used (self-evaluation feedback) would be beneficial for parents.

Comments regarding self-evaluation feedback:

_________________________________________________________________________________________
Appendix I

Parent Interview Questions

Parent Name: ___________________________________________  Date: __________________________

1. Do you feel that the HELPS Program helped your child? How?

________________________________________________________________________

________________________________________________________________________

2. Did your child enjoy the HELPS Program? Why or why not?

________________________________________________________________________

________________________________________________________________________

3. Did you enjoy using the HELPS Program with your child? Why or why not?

________________________________________________________________________

________________________________________________________________________

4. What worked with the HELPS Program?

________________________________________________________________________

________________________________________________________________________

5. What parts of the HELPS Program were challenging to implement?

________________________________________________________________________

________________________________________________________________________

6. How would you change the HELPS Program?

________________________________________________________________________

________________________________________________________________________

7. Did the weekly observation and feedback meetings help your implementation of the HELPS Program?

________________________________________________________________________

________________________________________________________________________
8. When in the Self-Evaluation phase, did completing the Observation Checklist help you implement the HELPS Program better? If so, how? If not, why not?

________________________________________________________________________

________________________________________________________________________

9. Was the amount of time required to evaluate yourself acceptable?

________________________________________________________________________

________________________________________________________________________

10. When in the verbal performance feedback phase, did the feedback provided to you help you implement the HELPS Program better? If so, how? If not, why not?

________________________________________________________________________

________________________________________________________________________

11. Was the amount of time required to receive the verbal feedback acceptable?

________________________________________________________________________

________________________________________________________________________

12. Was the self-evaluation or the verbal feedback more helpful to you? Why?

________________________________________________________________________

________________________________________________________________________

13. What other information would you like to share?

________________________________________________________________________

________________________________________________________________________
Appendix J

Program Implementation Quiz

True/False

Indicate whether the sentence or statement is true or false.

_____ 1. During the Phrase Drill portion of the intervention, it is okay to say, “Repeat after me” instead of, “Read after me.”

_____ 2. If a child reaches his goal on the first reading, you immediately chart his performance.

_____ 3. When graphing, you should circle the point that signifies the beginning of each session.

_____ 4. When graphing, you should connect the points of different passage numbers (e.g. passages 6, 7, and 8).

_____ 5. During the comprehension portion of a session, you should let the child continue to view the story.

_____ 6. You should be genuine when giving praise and try to avoid using the same praise statements repeatedly.

_____ 7. If a child does not meet his goal, the next step is for you to do model reading for the student.

_____ 8. You should let the child read to the end of the story each time he reads.

_____ 9. “Tell me what happened in the story” is a sufficient example of what to say for the instructions to begin the comprehension component.

_____ 10. For the first reading of the day, the child’s one-minute score should be recorded, but depending on the child’s performance, he may be allowed to read for up to 90 seconds.

_____ 11. A child can meet a “goal” twice per session.
12. As you give directions for the child to begin reading the first story of the day, he or she should not be able to view the beginning of the story.

**Multiple Choice**

13. If child does not meet her goal at the beginning of the session, but puts forth good effort during the rest of the session, the child earns
   a. two stars
   b. one star
   c. an automatic drawing from the prize bag
   d. no stars

14. Up to how many errors can a child make in order to still meet her goal?
   a. 0
   b. 1
   c. 2
   d. 3

15. If a child does not reach the 100 WCPM criterion on the first passage you
   a. don’t have to do the retell portion.
   b. ask her to immediately reread the passage
   c. still complete the retell portion.
   d. go directly to the phrase drill phase.

16. If a child reads the entire passage in less than one minute, you
   a. pretend that she read the passage in exactly one minute.
   b. write down the time that it took her to complete the passage and record the number of WCPM and WIPM during the reading on the tracking form.
   c. write down the time on the tracking form.
   d. give her two stars for her performance.
17. To keep track of what the one-minute bracket represents, you should
   a. record a 1, 2, or 3 with each bracket.
   b. not do anything; it doesn’t matter what order the scores are recorded.
   c. circle the bracket from reading 3.
   d. only put a bracket after the last word of the first reading.

18. If a child reads the passage with no errors, during a phrase drill
   a. you should skip the phrase drill procedure.
   b. you should spend time defining word that the child may not know.
   c. you should select at least 2 difficult portions of the passage to review using phrase
      drill procedures.
   d. randomly select words to practice using phrase drill procedures.

19. During the Model Reading phase, when reading the passage to the child, you should
   a. read just a little slower than the child reads.
   b. read at a pace that is just a little faster than the child’s reading ability.
   c. read the story with proper expression.
   d. both B and C.

20. If the child meets her goal, the first thing you should do with Passage B is
   a. phrase drill.
   b. have the child read it.
   c. read the passage to the student.
   d. ask the child what he/she knows about the story

21. If the child lands on a dark square, she
   a. picks a prize from the prize box.
   b. has to do an additional reading.
c. draws a ticket from the bonus bag.
d. answers what she knows about the story.

22. When graphing performance, you should graph
   a. the first reading.
   b. the second reading.
   c. the third reading.
   d. A and C

23. If a child meets her goal, what is the correct order of the protocol for Passage B?
   a. Reading 1; Phrase-drill; Reading 2; Retell Check; Reading 3
   b. Phrase-drill; Reading 1; Model Reading; Reading 2
   c. Reading 1; Phrase-drill; Reading 2; Model Reading; Reading 3
   d. Reading 1; Reading 2; Phrase-drill; Model Reading; Reading 3

**Short Answer**

25. If child does not meet her goal during the first reading, list what you would do for the rest of the session (include all steps up to giving the student stars and recording data on the tracking form):

26. If a child reads 78 words, and make 4 mistakes; write the number of WCPM and WIPM
   WCPM: ___________ WIPM: ___________

27. Please sketch a graph to report these data.
   9/1/07
   Passage 16
   Reading 1: WCPM = 78, WIPM = 3
   Reading 2: WCPM = 84, WIPM = 2
   Reading 3: WCPM = 95, WIPM = 1
9/2/07
Passage 16
Reading 1: WCPM = 95, WIPM = 2
Reading 2: WCPM = 100, WIPM = 4
Reading 3: WCPM = 115, WIPM = 2
9/3/07
Passage 16
Reading 1: WCPM = 105, WIPM = 1
Passage 17
Reading 1: WCPM = 87, WIPM = 4
Reading 2: WCPM = 94, WIPM = 3
Reading 3: WCPM = 98, WIPM = 4
9/4/07
Passage 17
Reading 1: WCPM = 99, WIPM = 3
Reading 2: WCPM = 108, WIPM = 2
Reading 3: WCPM = 115, WIPM = 2
Appendix K

HELPS One-on-One Program: Scripted Directions

Introductory statement & expectations (includes Verbal Cuing Procedure):
1. Say to the child, <Child’s Name>, 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 read to you). I also want you to remember the difficult words that we practice.

Directions to administer before a Timed Reading (as part of the Repeated Reading Procedure):
1. Place the teacher copy of the reading passage in front of you but shielded so your child can’t see what you record.
2. Place the student copy of the reading passage in front of your child, but cover the passage until you are ready for Step 4 below. (Do this so the student does not begin reading while you provide directions).
3. Say to your child, “Here is a story that I would like for 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 do your best reading.”
4. Say, “Begin!” and start your stopwatch when your child says the first word.
5. Mark your child’s incorrect words according to the Timed Reading Scoring Rules.
6. At the end of one minute, place a closed bracket after the last word.
7. If your child reads so fast that no expression is given, remind him/her that when he/she reads the next story, you want him/her to read at a comfortable rate (i.e., with good expression, like when you read).
8. Remove both copies of the reading passage.

Directions for administering the Retell Check Procedure:
1. Remove the student passage in a way to ensure your child cannot review the passage during the Retell Check.
2. Say to your child, “Now I want you to tell me everything you remember about the story you just read. Try to tell me what happened in the correct order.”
3. Start your stopwatch and stop the retell check in about 30-45 seconds. Use prompts or follow-up questions as appropriate.
4. If your child clearly struggles to remember parts of the story during his/her retell, note this on your child’s tracking sheet and use this information when determining whether your child met his/her reading goal.

Directions for administering Phrase-Drill Error Correction Procedure:
1. Say to your child, “Now we are going to practice some of the words you missed.”
2. Point to the first error word, say the word, and then say, “Read this after I do. <Read the 2-8 word phrase containing the error word>. Again, Again.” In essence, allow your child to read the phrase 3 times. Make sure your child points to the words being read; children will sometimes just “memorize” the phrase and repeat it.
3. Repeat the above procedures for all unique error words in the passage (up to 5 or until time permits).
i. If your child makes 1 or fewer errors, practice 1-3 phrases your child read less fluently. Use the same procedures as above, except say, “Now we are going to practice some words you read correctly, but they are difficult and we should practice them.”

4. Praise your child for every two to three sets of phrase-drills.

**Directions for parent to read passage aloud (Modeling Procedure):**

1. Say to your child, “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. When I stop, 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 doing your best.”

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 your child is following along with his/her finger and prompt him/her to do this, if necessary.

3. While reading the passage, stop 5-7 times in order to have your child read the word that immediately follows the word you just read.

4. At the end of the activity, praise your child for his/her effort.
Appendix L

HELPs One-on-One Program: Tips and Reminders for Implementation

General Implementation Procedure
☐ Parent 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, Star Chart, Bonus Bag, Implementation Flow Chart, and Scripted Directions. Also, the prize box was reasonably accessible.
☐ Parent used Scripted Directions or Abbreviated Directions as advised at top of Abbreviated Directions.

Repeated Reading (and Timed Reading) Procedure
☐ After each student reading, parent indicated on examiner passage (w/ a bracket) the # of words read in one min. If parent recorded all student readings with dry-erase marker before transferring scores to Progress Tracking Form, he:
☐ Put the appropriate number (i.e., 1, 2, or 3) next to the one-minute bracket.
☐ Marked student errors differently during each reading (e.g., first reading = slash, second reading = underline, third reading = circle).

Retell Check Procedure
☐ Before prompting student to begin the Retell Check, parent made sure student could not review the passage during the Retell Check.
☐ Parent used broad follow-up questions to solicit student’s retell only if student was unable to retell the passage for approximately 30 seconds.
☐ Parent implemented Retell Check for no more than 45 seconds unless he/she made a decision prior to the session to lengthen the Retell Check.

Goal Setting Procedure
☐ Parent told student he/she met the Reading Goal.

Phrase-Drill Error Correction Procedure
☐ Parent asked student to practice “logical” phrases.
☐ Parent told student to “READ” the phrases, and did not ask the student to “SAY” or “REPEAT” phrases.
☐ Parent had student practice all incorrectly read words (up to 5 or until time permits)
☐ Parent pointed (or had the student point) to each word practiced.
☐ If student practiced words that were read correctly but less fluently (see above step), parent explained to student that he/she read the words correctly, but will practice them because they are difficult.
☐ If student made 1 or fewer errors, the parent told the student to practice 1-3 words or phrases that were read less fluently.

Modeling Procedure
☐ Parent read aloud at a pace just a little faster than the student’s reading ability.
☐ Parent read with good expression.
☐ Parent read at a volume the student could clearly hear.
☐ Parent paused 5-7 times to have student read the next word in the passage.

Performance Feedback (Graphing) Procedure
☐ While graphing, parent gave verbal feedback and praise regarding the student’s WCPM and WIPM scores.
☐ Parent graphed WCPM and WIPM on 2 or 3 readings (3 if the Goal was met; 2 if the Goal was not met).
☐ Parent connected lines between WCPM (and WIPM) scores only for scores of the same passage.
☐ Parent circled the data point and session number when the student began a new passage.
**Motivational (Reward) Procedure**

- Throughout the session, parent provided a minimum of three different praise statements regarding the student's reading behavior.
- When awarding stars on Star Chart, parent accurately told student why he/she earned each star.
- With enthusiasm, parent 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 student for specific reading behaviors or improvements at the end of the session.
- If the student landed on OR passed a shaded square on Star Chart, student was allowed to select a ticket from the bonus bag and parent correctly recorded the bonus stars written on the ticket.
- Parent conveyed that improved reading skills, rather than the opportunity to earn stars/prizes, is the primary reason the student should put forth effort during each HELPS session.

**Using the Progress Tracking Form**

- After finishing session, parent completed Progress Tracking Form before erasing data from examiner passage.
- Parent recorded 2 or 3 sets of WCPM/WIPM scores on the Progress Tracking Form, as determined by whether the student met his/her Reading Goal on passage A (3 sets of scores were recorded if Goal was met; 2 sets of scores were recorded if Goal was not met).
- Parent recorded relevant information in the Notes column of the Progress Tracking Form (e.g., student difficulties with Retell Check, behavior problems, attention difficulties, etc.).

**Total steps applicable = 31 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 the Student’s Timed Readings**

Discrepancies / Total words read:

Reading 1: ____/ ____ ISRA%:_____  Reading 2: ____/ ____ ISRA%:_____

Reading 3: ____/ ____ ISRA%:_____  Reading 4: ____/ ____ ISRA%:_____

(When applicable)
Appendix M

HELPS One-on-One Program: Progress Tracking Form

Student Name: ______________________________  Grade: 2  Parent Name: ______________________________

<table>
<thead>
<tr>
<th>Session #</th>
<th>Date</th>
<th>1st story read</th>
<th>100 WCPM on 1st reading of passage A? (Y or N)</th>
<th>Student passes retell check? (Y or N)</th>
<th>WCPM/WIPM Timed Reading #1</th>
<th>WCPM/WIPM Timed Reading #3</th>
<th>WCPM/WIPM Timed Reading #1</th>
<th>WCPM/WIPM Timed Reading #3</th>
<th>Last story read</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix N
Appendix O

Reading with Effort Star Chart

(15 Stars = 1 prize from the special prize box)

Student Name: ________________________________

<p>| | | | | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix P

HELPS One-on-One Program Implementation Flow Chart

1. Parent reads introductory statements and expectations

2. Student Timed Reading (TR) with Passage A

3. Retell Check

   Student meets reading goal
   Student does not meet reading goal

   (See table below for goals according to the student’s grade level)

4a. Deliver Praise & Graph Passage A

5a. Student TR – Passage B, 1st time

6a. Phrase-drill procedure

7a. Student TR – Passage B, 2nd time

8a. Modeling Procedure

9a. Student TR – Passage B, 3rd time

4b. Modeling procedure

5b. Student TR – Passage A, 2nd time

6b. Phrase-drill procedure

7b. Student TR – Passage A, 3rd time

8b. Phrase-drill procedure

Graph 1st and 3rd TR of Passage B and provide praise and feedback

Award stars on Star Chart

Record student data on Progress Tracking Form

Graph 1st and 3rd TR of Passage A and provide praise and feedback

Award stars on Star Chart

Record student data on Progress Tracking Form

Reading Goals According to the Student’s Grade Level

<table>
<thead>
<tr>
<th>Grade</th>
<th>WCPM with Passage A</th>
<th>WIPM with Passage A</th>
<th>Retell Check with Passage A</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Grade</td>
<td>80 or more</td>
<td>3 or less</td>
<td>Adequately retells story*</td>
</tr>
<tr>
<td>Second Grade</td>
<td>100 or more</td>
<td>3 or less</td>
<td>Adequately retells story*</td>
</tr>
<tr>
<td>Third Grade</td>
<td>120 or more</td>
<td>3 or less</td>
<td>Adequately retells story*</td>
</tr>
<tr>
<td>Fourth Grade</td>
<td>135 or more</td>
<td>3 or less</td>
<td>Adequately retells story*</td>
</tr>
</tbody>
</table>
* For example, student retells parts of the story for at least 30 seconds or otherwise correctly states names of characters and major events in the story. Retell of the story in the correct sequential order of major events is encouraged but not required to pass the Retell Check.
Appendix Q

Steps and Guidelines for Observing Parents During Their Implementation of the HELPS One-on-One Program: Baseline Phase

- 1. On the Observation Summary Form (OSF), record (a) your name, (b) the parent’s name, (c) the date, (d) the observation (OBS) session number (e.g., if the parent has been observed 4 times previously, write “5” because this is the 5th OBS session), (e) the name of student receiving HELPS, and (f) the student’s HELPS session number.

- 2. Wait patiently until the parent is ready to begin implementing HELPS with the student and observe the parent’s organization and preparation for the session.

- 3. **When the parent begins Step 1 of HELPS implementation, start your stopwatch** to begin monitoring the total time it takes the parent to complete the session.

- 4. Throughout the HELPS implementation session, record all steps the parent completes correctly on the Observation Checklist for Implementing the HELPS Program. Steps should be recorded for implementation of both Core Procedures and Tips and Reminders.

- 5. For each Timed Reading the student completes during the HELPS session, follow along on your examiner copy and record student errors. You do not need to time the student, but make sure the parent accurately times the student for one minute.

- 6. At the end of the entire HELPS session (after the parent completes the Progress Tracking form), **stop your stopwatch and record the following information on the OSF:** (a) whether the student met his/her goal, (b) the duration of the HELPS session in minutes and seconds, (c) the number or % of steps the parent 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 parent completed accurately from the Tips/Reminders checklist, (f) your evaluation of the parent’s enthusiasm during the session, and (g) your evaluation of the parent’s organization during the session.

- 7. On your Observation Checklist, record all Inter-Scorer Reliability Agreement (ISRA) data.

- 8. **Reset and start stopwatch again.** Ask the parent if he/she has any questions about HELPS implementation procedures (or the program, in general) and record those questions as needed. **Remember, no feedback.**

- 9. Thank the parent for his/her time and effort and conclude the OBS session. **Overall, the parent should finish each OBS session feeling positive and assured that, although she/he is not receiving any feedback, we are using our observation data to better understand how to support parents.**

- 10. Stop your stopwatch and record on your OSF: (a) the duration of the post-session OBS meeting, (b) whether all parent questions were recorded, and (c) any additional, meaningful notes about the meeting.

- 11. Review steps 1-10 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 parent.

Percentage of steps completed = total steps completed / total steps applicable: ___/___ x 100 = _____%
Appendix R

Steps and Guidelines for Observing Parents During Their Implementation of the HELPS One-on-One Program: Self-Eval Phase

☐ 1. On the Observation Summary Form (OSF), record (a) your name, (b) the parent's name, (c) the date, (d) the observation (OBS) session number (e.g., if the parent has been observed 4 times previously, write “5” because this is the 5th OBS session), (e) the name of student receiving HELPS, and (f) the student's HELPS session number.

☐ 2. Wait patiently until the parent is ready to begin implementing HELPS with the student and observe the parent's organization and preparation for the session.

☐ 3. **When the parent begins Step 1 of HELPS implementation, start your stopwatch** to begin monitoring the total time it takes the parent to complete the session.

☐ 4. Throughout the HELPS implementation session, record all steps the parent completes correctly on the Observation Checklist for Implementing the HELPS Program. Steps should be recorded for implementation of both Core Procedures and Tips and Reminders.

☐ 5. For each Timed Reading the student completes during the HELPS session, follow along on your examiner copy and record student errors. You do not need to time the student, but make sure the parent accurately times the student for one minute.

☐ 6. At the end of the entire HELPS session (after the parent completes the Progress Tracking form), stop your stopwatch and record the following information on the OSF: (a) whether the student met his/her goal, (b) the duration of the HELPS session in minutes and seconds, (c) the number or % of steps the parent 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 parent completed accurately from the Tips/Reminders checklist, (f) your evaluation of the parent's enthusiasm during the session, and (g) your evaluation of the parent's organization during the session.

☐ 7. On your Observation Checklist, record all Inter-Scorer Reliability Agreement (ISRA) data.

☐ 8. **Reset and start your stopwatch again.** Next, have the parent (or assist the parent) rewind the audiotape to the beginning of the current session. Give the parent a copy of the Parent’s Observation Checklist for Implementing the HELPS One-on-One Program to complete while listening to the entire recorded session.

☐ 9. After the parent finishes listening to the recorded session and completing the form, provide a few minutes for the parent to review the form.

☐ 10. Ask the parent if he/she has any questions about HELPS implementation procedures (or the program, in general) and record those questions as needed. **Remember, no feedback.**

☐ 11. Thank the parent for his/her time and effort and conclude the OBS session. **Overall, the parent should finish each OBS session feeling positive and assured that, although she/he is not receiving any feedback, we are using our observation data to better understand how to support parents.**
12. Stop your stopwatch and record on your OSF: (a) the duration of the post-session OBS meeting, (b) whether all parent questions were recorded, (c) all Core Procedures not implemented, and (d) any additional, meaningful notes about the meeting.

13. Review steps 1-12 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 parent.

Percentage of steps completed = total steps completed / total steps applicable: ____/____ x 100 = ____%
Appendix S

Parent's Observation Checklist for Implementing the HELPS One-on-One Program

Parent name: ______________________________  Child’s Name: ______________________________

Observer: __________________  Date: __________________

Implementation of Core Procedures

(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)

List of Steps (if Goal is met): 1 ☐; 2 ☐; 3 ☐; 4a ☐; 5a ○; 6a ○; 7a ○; 8a ○; 9a ○; 10a ○; 11a ○; 12a ○

List of Steps (if Goal is not met): 1 ○; 2 ○; 3 ○; 4b ○; 5b ○; 6b ○; 7b ○; 8b ○; 9b ○; 10b ○; 11b ○

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

Implementation of Tips and Reminders (steps missed below are considered minor errors)

(Place “✓” in the box for completed steps; Place “X” in the box for skipped steps; Place circle “O” around the box for non-applicable [NA] steps)

General Implementation Procedure

☐ I 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, Star Chart, Bonus Bag, Implementation Flow Chart, and Scripted Directions. Also, the prize box was reasonably accessible.

☐ I used Scripted Directions or Abbreviated Directions as advised at top of Abbreviated Directions.

Repeated Reading (and Timed Reading) Procedure

☐ After each student reading, I indicated on examiner passage (w/a bracket) the # of words read in one min

If I recorded all student readings with dry-erase marker before transferring scores to Progress Tracking Form, I:

☐ Put the appropriate number (i.e., 1, 2, or 3) next to the one-minute bracket.

☐ Marked student errors differently during each reading (e.g., first reading = slash, second reading = underline, third reading = circle).

Retell Check Procedure

☐ Before prompting my child to begin the Retell Check, I made sure my child could not review the passage during the Retell Check.

☐ I used broad follow-up questions to solicit my child’s retell only if my child was unable to retell the passage for approximately 30 seconds.

☐ I implemented the Retell Check for no more than 45 seconds unless I made a decision prior to the session to lengthen the Retell Check.

Goal Setting Procedure

☐ I told my child he/she met the Reading Goal.

Phrase-Drill Error Correction Procedure

☐ I asked my child to practice “logical” phrases.

☐ I told my child to “READ” the phrases, and did not ask my child to “SAY” or “REPEAT” phrases.

☐ I had my child practice all incorrectly read words (up to 5 or until time permits)
I pointed (or had my child point) to each word practiced.
If my child practiced words that were read correctly but less fluently (see above step), I explained to my child that he/she read the words correctly, but will practice them because they are difficult.
If my child made 1 or fewer errors, I told my child to practice 1-3 words or phrases that were read less fluently.

**Modeling Procedure**
- I read aloud at a pace just a little faster than my child’s reading ability.
- I read with good expression.
- I read at a volume my child could clearly hear.
- I paused 5-7 times to have my child read the next word in the passage.

**Performance Feedback (Graphing) Procedure**
- While graphing, I gave verbal feedback and praise regarding my child’s Words Correct Per Minute (WCPM) and Words Incorrect Per Minute (WIPM) scores.
- I graphed WCPM and WIPM on 2 or 3 readings (3 if the Goal was met; 2 if the Goal was not met).
- I connected lines between WCPM (and WIPM) scores only for scores of the same passage.
- I circled the data point and session number when my child began a new passage.

**Motivational (Reward) Procedure**
- Throughout the session, I provided a minimum of three different praise statements regarding my child’s reading behavior.
- When awarding stars on Star Chart, I accurately told my child why he/she earned each star.
- With enthusiasm, I 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 my child for specific reading behaviors or improvements at the end of the session.
- If the my child landed on OR passed a shaded square on Star Chart, my child was allowed to select a ticket from the bonus bag and I correctly recorded the bonus stars written on the ticket.
- I conveyed that improved reading skills, rather than the opportunity to earn stars/prizes, is the primary reason my child should put forth effort during each HELPS session.

**Using the Progress Tracking Form**
- After finishing session, I completed the Progress Tracking Form before erasing data from examiner passage.
- I recorded 2 or 3 sets of WCPM/WIPM scores on the Progress Tracking Form, as determined by whether my child met his/her Reading Goal on passage A (3 sets of scores were recorded if Goal was met; 2 sets of scores were recorded if Goal was not met).
- I recorded relevant information in the Notes column of the Progress Tracking Form (e.g., difficulties with Retell Check, behavior problems, attention difficulties, etc.).

Total steps applicable = 31 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: _______%
Appendix T

Steps and Guidelines for Observing Parents During Their Implementation of the HELPS One-on-One Program: PF Phase

☐ 1. On the Observation Summary Form (OSF), record (a) your name, (b) the parent’s name, (c) the date, (d) the observation (OBS) session number (e.g., if the parent has been observed 4 times previously, write “5” because this is the 5th OBS session), (e) the name of student receiving HELPS, and (f) the student’s HELPS session number.

☐ 2. Wait patiently until the parent is ready to begin implementing HELPS with the student and observe the parent’s organization and preparation for the session.

☐ 3. **When the parent begins Step 1 of HELPS implementation, start your stopwatch** to begin monitoring the total time it takes the parent to complete the session.

☐ 4. Throughout the HELPS implementation session, if the parent implements a step out of order, forgets to implement a step, or makes a major procedural error when implementing a particular step (and the parent does not self-correct the mistake within 7-10 seconds), use immediate corrective feedback regarding the error made. Be sure to correct the mistake in a respectful, clear, and concise way. The table at the end of this checklist specifies all major procedural errors. Steps missed on the Tips and Reminders checklist are considered minor errors and will be discussed with the parent after he/she completes the session (as described below).

☐ 5. Throughout the HELPS implementation session, record all steps the parent completes correctly on the Observation Checklist for Implementing the HELPS Program. Steps should be recorded for implementation of both Core Procedures and Tips and Reminders.

☐ 6. For each Timed Reading the student completes during the HELPS session, follow along on your examiner copy and record student errors. You do not need to time the student, but make sure the parent accurately times the student for one minute.

☐ 7. At the end of the entire HELPS session (after the parent completes the Progress Tracking form), stop your stopwatch and record the following information on the OSF: (a) whether the student met his/her goal, (b) the duration of the HELPS session in minutes and seconds, (c) the number or % of steps the parent 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 parent completed accurately from the Tips/Reminders checklist, (f) your evaluation of the parent’s enthusiasm during the session, and (g) your evaluation of the parent’s organization during the session.

☐ 8. On your Observation Checklist, record all Inter-Scorer Reliability Agreement (ISRA) data.

☐ 9. **Reset and start your stopwatch again.** Next, identify at least 2-3 steps that the parent 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 parent made since an earlier OBS session (related to step 17 below).

☐ 10. Share the data you recorded in parts b, c, d, e, f, and g (from step 7 above) with the parent.

☐ 11. If you provided immediate feedback during the HELPS session (i.e., major implementation steps were skipped or implemented incorrectly), briefly review those implementation errors and ask the parent if he/she has any questions about those steps. When appropriate, provide a rationale for why a step should be performed in a particular way. (If you are uncertain about this information, write down the question, consult the HELPS Teacher’s Manual, and later share that information with the parent).

☐ 12. If there were steps from the Tips and Reminders checklist the parent did not implement, review those missed steps with the parent and, if needed, discuss any questions the teach has about those missed steps.

☐ 13. If you did not rate the parent’s enthusiasm and/or organization as “outstanding,” provide a rationale of your evaluation to the parent and discuss the situation as needed.

☐ 14. If applicable, provide additional feedback (i.e., feedback not related to the Implementation Protocol or Tips and Reminders Checklist) to the parent that will likely help him/her implement HELPS better in the future. For example, you may offer advice about how to organize HELPS materials in the most effective and time efficient way.

☐ 15. Ask the parent if he/she has any questions about HELPS implementation procedures (or the program, in general) and answer/discuss those questions as needed.

☐ 16. Record the topics discussed in steps 14 and/or 15 on your OSF.

☐ 17. At the end of the OBS session, identify 1-3 things (as deemed appropriate/applicable) the parent should improve upon during subsequent HELPS sessions with students (these are considered “targeted steps” for the parent to improve). Make sure the parent 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.

☐ 18. Thank the parent for his/her time and effort and conclude the OBS session. Overall, the parent should finish each OBS session feeling positive and better prepared to implement the HELPS Program, rather than feeling judged or deemed inadequate.

☐ 19. Stop your stopwatch and record on your OSF:

(a) the duration of the post-session OBS meeting,
(b) whether all parent 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 17 above), specify these targeted steps on your OSF or write “None” if no Core Procedures or Tips/Reminders were missed.
20. Review steps 1-19 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 parent.

Percentage of steps completed = total steps completed / total steps applicable: ____/____ x 100 = ____%