4-29-2013

Direct Training to Increase Inter-rater Agreement between an Observer’s and Teachers’ Self-Report Ratings of Treatment Integrity

Lindsay M. Fallon PhD
lindsay.m.fallon@gmail.com

Follow this and additional works at: http://digitalcommons.uconn.edu/dissertations

Recommended Citation
Fallon, Lindsay M. PhD, "Direct Training to Increase Inter-rater Agreement between an Observer’s and Teachers’ Self-Report Ratings of Treatment Integrity" (2013). Doctoral Dissertations. 49.
http://digitalcommons.uconn.edu/dissertations/49
4-29-2013

Direct Training to Increase Inter-rater Agreement between an Observer’s and Teachers’ Self-Report Ratings of Treatment Integrity

Lindsay M. Fallon PhD

Follow this and additional works at: http://digitalcommons.uconn.edu/dissertations
Measuring an implementer’s treatment integrity, specifically an implementer’s adherence to steps of an intervention, can be done via direct (e.g., observation) or indirect (e.g., self-report) methods of assessment. Direct observation is a widely used and accepted method of data collection in research due to its technical adequacy. However, direct observation is resource intensive, often making it impractical outside of research. Self-report measures of adherence can be less resource intensive and are commonly used in school settings, yet results from previous research indicate that implementers frequently overestimate their adherence when using self-report measures. To address this issue, results from research that build support for teacher self-report as a reliable method of treatment integrity assessment are needed. As such, the objective of the current study was to improve inter-rater agreement (IRA) between teachers’ adherence self-report ratings and ratings provided by an observer. The student investigator (i.e., primary observer) observed instructional practice during baseline. Then, after a brief indirect training on the intervention, the primary observer and teachers rated teacher adherence to an explicit instruction intervention. When it was determined that the teachers’ adherence ratings did not adequately agree with the observer’s, teachers were staggered into a phase in which they received direct training on the intervention steps to assess if a change in IRA occurred. Results indicate that after direct intervention training, IRA between the primary observer and teachers improved.
Direct Training to Increase Inter-rater Agreement between an 
Observer’s and Teachers’ Self-Report Ratings of Treatment Integrity

Lindsay Marie Fallon

B.A., University of Wisconsin-Madison, 2004
M.S.Ed., Queens College, City University of New York, 2007
M.A., University of Connecticut, 2009

A Dissertation
Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy
at the
University of Connecticut
2013
Copyright by

Lindsay Marie Fallon

2013
Approval Page

Doctor of Philosophy Dissertation

Direct Training to Increase Inter-rater Agreement between an Observer’s and Teachers’ Self-Report Ratings of Treatment Integrity

Presented by
Lindsay Marie Fallon, B.A., M.S.Ed., M.A.

Major Advisor ________________________________________________
Lisa M. H. Sanetti, Ph.D.

Associate Advisor _____________________________________________
Sandra M. Chafouleas, Ph.D.

Associate Advisor _____________________________________________
Michael N. Faggella-Luby, Ph.D.

University of Connecticut
2013
Acknowledgements

Looking back, I am amazed at the number of individuals who made this feat possible by lending me their time and ideas. I express my sincerest gratitude to the faculty of the program in school psychology and special education at the University of Connecticut, particularly members of my doctoral committee and most notably my advisor, Lisa. Your encouragement pushed me to achieve more than I could have imagined five years ago and I am inspired by your ability to manage so many personal and professional responsibilities in a manner that appears effortless. I must also thank my classmates. Through your thoughtful consideration of my ideas during countless hours of discourse, I gained the motivation I needed to push through this process. I also gained several dear friends. To my parents, sisters and close friends: your unconditional love and encouragement were integral to my perseverance in pursuing this project and degree. Thank you. To my in-laws who generously dropped everything to watch Zach so that I may collect data on the many chilly mornings that our babysitter canceled: thank you. I am so appreciative of your support. To Zach for being a great inspiration and always making me smile. I love you very much. But most of all, to my loving, nurturing, hilarious husband, Derek, who helped to make this project and degree possible in more ways than I can count. I dedicate this dissertation to you.
TABLE OF CONTENTS

Approval Page ................................................................. iii
Acknowledgements .............................................................. iv
Table of Contents .............................................................. v
List of Tables ................................................................. vi
List of Figures ................................................................. vii
List of Appendices ........................................................... viii

CHAPTER I: INTRODUCTION

Statement of Problem ...................................................... 1
Purpose of Study ............................................................. 3

CHAPTER II: REVIEW OF THE LITERATURE

Treatment Integrity Assessment ........................................... 5
Determining Agreement between Raters............................... 9
Teacher Training ............................................................. 11
Explicit Instruction ........................................................ 14
Research Question and Hypothesis ....................................... 17

CHAPTER III: METHOD

Participants ................................................................. 19
Instrumentation ............................................................ 20
Dependent Variable Definitions ........................................... 23
Design ................................................................................. 24
Procedures ........................................................................... 25
Data Analysis ....................................................................... 35

CHAPTER IV: RESULTS ......................................................... 38

CHAPTER V: DISCUSSION ..................................................... 47

REFERENCES ...................................................................... 56

TABLES .............................................................................. 72

FIGURES ............................................................................ 78

APPENDICES ...................................................................... 85
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>Effect Size Calculations</td>
<td>73</td>
</tr>
<tr>
<td>Table 2</td>
<td>Cohen’s Weighted Kappa Coefficients</td>
<td>74</td>
</tr>
<tr>
<td>Table 3</td>
<td>Teachers’ Average Adherence to Steps of Explicit Instruction Stages by Intervention Phase</td>
<td>75</td>
</tr>
<tr>
<td>Table 4</td>
<td>Inter-observer Agreement for Primary and Secondary Observers’ Rating</td>
<td>76</td>
</tr>
<tr>
<td>Table 5</td>
<td>Results from Usage Rating Profile-Intervention</td>
<td>77</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Figure 1</td>
<td>Ratings on DBR Scale for Advance Organizer Stage</td>
<td>79</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Ratings on DBR Scale for Student Enlistment Stage</td>
<td>80</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Ratings on DBR Scale for Post-Organizer Stage</td>
<td>81</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Difference Between Teachers’ and Observer’s Ratings on DBR Scale for Advance Organizer Stage</td>
<td>82</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Difference Between Teachers’ and Observer’s Ratings on DBR Scale for Student Enlistment Stage</td>
<td>83</td>
</tr>
<tr>
<td>Figure 6</td>
<td>Difference Between Teachers’ and Observer’s Ratings on DBR Scale for Post-Organizer Stage</td>
<td>84</td>
</tr>
</tbody>
</table>
LIST OF APPENDICES

Appendix A  Decision Tree for Determining Adequate Agreement .................. 86
Appendix B  Stages of the Explicit Instruction Intervention .......................... 87
Appendix C  Expanded Table of Explicit Instruction Definitions and Sample Dialogue .................................................. 88
Appendix D  Teacher Daily Instruction Form .................................................. 92
Appendix E  Social Validity Measure ................................................................. 93
Appendix F  Teacher Information Sheet and Consent Form ............................ 95
Appendix G  Abbreviated Table of Explicit Instruction Definitions ................. 98
Appendix H  Expert Raters for Primary Observer Training ............................ 99
Appendix I  Draft of Indirect Training Script .................................................. 100
Appendix J  Draft of Direct Intervention Training Script ............................... 103
Appendix K  Draft of Direct Intervention Training Slides for Video Presentation 105
Appendix L  Draft of Direct Assessment Training Script ............................... 111
Appendix M  Draft of Direct Assessment Training Slides for Video Presentation 113
Appendix N  Solutions for Deviations from Study Procedures ........................ 123
Chapter I: Introduction

Statement of the Problem

Over the past decade, legislation (i.e., Individuals with Disabilities Education Act [IDEA], 2004) has encouraged many schools to adopt a problem-solving framework such as response to intervention (RTI) to identify students in need of additional supports (e.g., academic, behavioral) to achieve successful outcomes. RTI involves delivering high-quality, empirically supported instruction and monitoring student progress (National Association of State Directors of Special Education, 2008). Students who do not respond as expected are provided additional intervention and their progress continues to be monitored. When delivering supports in a multi-tiered model, like RTI, student responsiveness to the provision of empirically supported interventions is the primary factor that determines the level of intervention required (Sanetti & Kratochwill, 2009a).

To determine if an intervention has an effect on student outcomes, data related to adult behavior (i.e., treatment integrity) are needed in addition to student progress monitoring data. That is, it is critical that when an intervention is implemented, treatment integrity data are collected and analyzed (Wilkinson, 2006). Broadly, treatment integrity is the degree to which interventions are implemented as planned (e.g., by teachers; Gresham, 1989).

The collection of treatment integrity data is important for evaluating an intervention’s effectiveness, particularly in school-based consultation (Noell, 2008). In school-based consultation, a consultant (e.g., a school psychologist) works with a consultee (e.g., teacher) to provide supports to a client (e.g., a student; behavioral consultation in the schools, Kratochwill & Bergan, 1990; school consultation, Erchul & Martens, 2010). Within a school-based consultation framework, services are indirect, the
goal is prevention and there is an emphasis on using a problem-solving approach to address students’ academic and behavioral concerns (Gansle & Noell, 2008). Being an indirect service delivery model, it is particularly critical in school-based consultation to collect and review treatment integrity data to determine if any noted changes in student outcomes are in fact due to the teachers’ implementation of a prescribed intervention (Noell, 2008). Without assessment of treatment integrity, school-based consultants cannot be certain that student outcomes, specifically poor student outcomes, are due to an ineffective intervention or an intervention that is poorly implemented (Sanetti & Kratochwill, 2009a).

Despite the importance of collecting treatment integrity data, particularly in school-based consultation, many school psychologists do not regularly collect data on teachers’ treatment integrity. Cochrane and Laux (2008) surveyed nationally certified school psychologists and found that although most respondents agreed that treatment integrity is important to evaluating intervention effectiveness, many also reported that treatment integrity data are not always collected in consultation with teachers. Among the most common reasons respondents gave for not assessing treatment integrity in consultation was a lack of time. Of the respondents who reported that treatment integrity is always (11.3%) or sometimes (41.6%) assessed in consultation with teachers, 36.9% indicated that treatment integrity data are collected through teacher self-report (versus 25.3% who employed direct observation). This may be because collecting teacher self-report data requires less time for consultants than conducting direct observation (Biggs, Vernberg, Twemlow, Fonagy, & Dill, 2008).

Although teacher self-report may be a more practical (i.e., less time-intensive) means of assessing treatment integrity data in school-based consultation, results from
research have demonstrated that many teachers (not all; Sanetti & Kratochwill, 2009b) overestimate their adherence to steps of an intervention when using self-report measures (e.g., Noell et al., 2005; Wickstrom, Jones, LaFleur, & Witt, 1998). Therefore, in treatment integrity and school-based consultation research, direct observation is frequently employed to assess treatment integrity as self-report data typically lack reliability (Gresham, 1989; Noell, 2008). This illustrates a discrepancy that exists between research and practice in the area of treatment integrity assessment.

**Purpose of the Study**

Development and empirical support for a feasible method of assessing treatment integrity is very much needed (Sanetti & Kratochwill, 2009a). As teacher self-report is often used by school psychologists to measure treatment integrity (Cochrane & Laux, 2008), and teachers prefer assessment methods that can be integrated into existing classroom routines versus those that might pose a diversion to instruction (Biggs et al., 2008), beginning to build support for the use of this method is perhaps the most logical place to start.

Because teachers often overestimate their adherence to steps of an intervention when compared to data collected via direct observation, one way to build support for teacher self-report is to increase inter-rater agreement (IRA) between an observer’s and teachers’ ratings of intervention implementation. IRA might be improved when teachers are provided with comprehensive, direct training on executing steps of the intervention or using a self-report measure rather than being given a brief, didactic (i.e., indirect) training, which is currently more common in school-based consultation (Sterling-Turner, Watson, & Moore, 2001). More comprehensive training on intervention implementation might be beneficial to curb teachers’ overestimation of adherence. Consultants often
underestimate the amount of training and support teachers require to implement evidence-based practices that target both improving academic and behavioral outcomes for students (e.g., explicit instruction; Erchul & Martens, 2010). Therefore, the purpose of this study was to evaluate if direct training will increase observed IRA between an observer’s and teachers’ self-report ratings of adherence to stages of explicit instruction from an indirect training phase. Results from the current study demonstrated an improvement in IRA after teachers received direct training on intervention steps. These findings could assist in building support for teacher self-report as a practical method of assessing treatment integrity that produces reliable data.
Chapter II: Review of the Literature

According to the Individuals with Disabilities Education Act of 2004, when evaluating a student’s eligibility to receive special education services, measures used in assessment must be technically sound, producing data that are reliable (as well as valid; IDEA, 2004). However, the law offers little guidance about the types of measures to use when collecting such data. IDEA (2004) also allows for some students to be identified as being in need of special education services through a RTI process, a problem-solving framework that has been widely adopted in the United States. Implementing a RTI framework requires (a) providing empirically supported instruction for all students, (b) monitoring student outcomes, (c) assessing the implementation of instruction, and (d) making data-based decisions about student responsiveness (Gresham, 2004). In other words, before making decisions about student responsiveness to an intervention, data are needed related to student performance as well as adult behavior (i.e., treatment integrity; Sanetti & Kratochwill, 2009a). Yet research is still emerging that identifies measures that are technically adequate as well as feasible to assist in these activities (Briesch, Chafouleas, & Riley-Tillman, 2010; Noell, 2008).

Treatment Integrity Assessment

In the last 30 years, studies that connect treatment integrity (Peterson, Homer, & Wonderlich, 1982) or treatment fidelity (Moncher & Prinz, 1991) to treatment outcomes have emerged in a variety of fields including mental health (Vermilyea, Barlow, & O’Brien, 1984), criminal justice (Henggeler, Melton, Brondino, Scherer, & Hanley, 1997) and, most relevantly, education (Greenwood, Terry, Arreaga-Mayer, & Finney, 1992; McEvoy, Shores, Wehby, Johnson, & Fox, 1990; Noell, Gresham, & Gansle, 2002; Noell, Witt, Gilbertson, Ranier, & Freeland, 1997). In studies of treatment integrity,
researchers have been interested in the degree to which an independent variable is manipulated as intended (Arkoosh et al., 2007), with the hypothesis that treatment integrity may serve as a moderator between intervention and outcome (Carroll et al., 2007). As summarized by Sanetti and Kratochwill (2009a), assessing treatment integrity might target measuring various aspects of intervention delivery: which steps of the intervention were delivered (e.g., adherence; Dane & Schneider, 1998), how much of the intervention was provided (e.g., dosage; Jones Clarke, & Power, 2008), how well the intervention was delivered (e.g., quality; Dane & Schneider, 1998), or the process by which intervention delivery occurred (e.g., exposure; Dane & Schneider, 1998).

In school-based research, direct and indirect methods of collecting treatment integrity data have been investigated. These methods have most commonly targeted the assessment of treatment adherence (Sanetti, Gritter, & Dobey, 2011). In many published studies, authors report having used direct methods such as direct observation (Burns, Peters, & Noell, 2008; DiGenarro, Martens, & Kleinmann, 2005; DiGenarro, Martens, & McIntyre, 2007). This is despite limitations to direct observation pertaining to content validity, observer bias, and participant reactivity (Gresham, Cook, Crews & Barreras, 2007). Direct observation is more often used for treatment integrity assessment in research because, as previously described, when researcher-collected and teacher self-report treatment integrity data have been compared, results suggest that many teachers overestimate their adherence to steps of an intervention (e.g., Noell et al., 2005; Wickstrom et al., 1998). Assuming direct observation provides a true score of implementation, this type of systematic overestimation on behalf of teachers is termed as fixed bias in classical measurement theory (Del Boca & Noll, 2000). It reflects a consistent distortion of implementation data that is not random, limiting the reliability of
data produced by teacher self-report. As a result, there is often low agreement between direct and indirect methods of treatment integrity assessment (Gresham, MacMillan, Beebe Frankenberger, & Bocian, 2000; Lane, 2007; Wickstrom et al., 1998).

Despite this lack of support for the technical adequacy of teacher self-report, many school psychologists rely on teachers to self-report treatment integrity data in schools (Cochrane & Laux, 2008). It is less resource-intensive than direct measures (Hersen, 2004; Riley-Tillman, Chafouleas, Briesch, & Eckert, 2008), may be viewed as less evaluative than direct observation (Biggs et al., 2008), and provides teachers with the opportunity to build new skills (i.e., in data collection), a goal targeted in school-based consultation (Erchul & Martens, 2010; Gresham, 1989). As such, research findings that support a method for improving the reliability of teacher self-report data are needed. To date, treatment integrity research that has assessed the reliability of teacher self-report has done so as secondary analysis; specifically researchers have typically reported average treatment integrity scores across methods (Noell et al., 2005; Wickstrom et al., 1998). Increasing the technical adequacy of more feasible, indirect methods of treatment integrity assessment (such as teacher self-report) has rarely been the primary focus of empirical study, but this research is needed (Sanetti & Kratochwill, 2009a).

**Teacher self-report.** If data produced by teacher self-report were more reliable, there are multiple ways in which teachers could self-report their classroom behavior. Teachers might describe their implementation via verbal self-report. This would involve a teacher indicating to a consultant whether intervention components were delivered as intended (Bramlett, Murphy, Johnson, Wallingsford, & Hall, 2002). Although this method is feasible, the possibility for social desirability bias might affect data collected via this method (Gresham et al., 2000) and it has yet to receive extensive empirical
review in the treatment integrity literature (e.g., Sanetti, Chafouleas, O’Keeffe, & Kilgus, 2011).

Teachers might also complete a treatment integrity checklist. A treatment integrity checklist provides a prompt to implementers to self-report the extent to which each component of a sequential list of intervention components was implemented as intended (Wilkinson, 2006). Although checklists can be flexible, quick and easy to complete (Wortham, 2008), they require interventions to be broken down into discrete steps and completed by teachers after implementation (Sanetti & Kratochwill, 2008). This may not be suitable for monitoring all interventions (e.g., an intervention with non-discrete steps or steps that occur in variable order).

Another self-report option is for teachers to rate their treatment integrity using a rating scale. To monitor treatment integrity over time, however, a rating scale should be appropriate for repeated use. A well-researched option for implementers to provide a brief, low-interference rating of relevant operationally-defined behaviors (e.g., implementation of intervention components) after a pre-specified observation period is Direct Behavior Rating (DBR; Christ, Riley-Tillman, & Chafouleas, 2009). DBR is a flexible, efficient tool that is designed to capture a rating in close proximity of a behavior and can be used repeatedly (Volpe & Briesch, 2012).

**Direct Behavior Rating.** DBR has been associated with positive outcomes when used for formative assessment (Riley-Tillman, Methe, & Weegar, 2009), progress monitoring (e.g., integrated into a Daily Behavior Report Card; Chafouleas, Riley-Tillman, & Sassu, 2006), behavioral screening (Chafouleas, Kilgus, & Hernandez, 2009), and home-school communication (Jaffery & Chafouleas, 2012). To date, many DBR evaluation studies have assessed the technical adequacy of rating three specific
behavioral targets: academic engagement, disruptive behavior, and respect (e.g., Chafouleas, Christ, & Riley-Tillman, 2009); however DBR can be used to assess a variety of behaviors. DBR can occur on a Single-item Scale (DBR-SIS), on which one behavior is targeted for rating (Chafouleas, Christ et al., 2009), or a Multi-item Scale (DBR-MIS). DBR-MIS is a scale in which multiple operationally defined target behaviors are typically rated on a unipolar graphic rating scale (Christ et al., 2009; Chafouleas, Riley-Tillman, & Christ, 2009). The scale consists of a fixed number of gradients with ordinal properties (e.g., 1 to 10) that are equitably spaced along a horizontal line (Chafouleas, Christ et al., 2009).

Although DBR scales can be created for a wide range of observable behaviors, with creation of a scale comes uncertainty about psychometric properties, specifically concurrent validity and treatment sensitivity (Volpe & Gadow, 2010). DBR is also limited somewhat by threats to internal validity associated with many self-report measures (e.g., overestimation, reactivity; Briesch et al., 2010). Nonetheless, the use of DBR is promising for the assessment of both child and adult behavior. It demands few resources, can be used daily by teachers to self-report behavior, is brief and adaptable to various behaviors of interest. Furthermore, ratings can also be easily compared for the purpose of calculating IRA.

**Determining Agreement between Raters**

**Inter-rater agreement.** IRA is calculated to determine if ratings provided by one rater are observed to be similar to the ratings provided by one or multiple additional raters (LeBreton, Burgess, Kaiser, Atchley, & James, 2003), specifically if scores provided by two or more raters are equivalent (LeBreton & Senter, 2008). IRA is often determined using percent agreement, Cohen’s kappa (Cohen, 1960), or Fleiss’ kappa
(Fleiss, 1971). Percent agreement is easy to calculate (as the number of times an agreement occurs divided by the number of times in which an agreement could occur multiplied by 100) but does not take into consideration agreement that occurs by chance. It also treats any difference between two ratings as the same: disagreement (Hunt, 1986). For example, disagreement is determined whether two raters rate a behavior within 2 points of one another or 5 points. This method does not account for the degree to which ratings differ, providing less information about rater behavior in treatment integrity data collection.

Cohen’s kappa, an alternative to percent agreement, takes into account agreement that occurs by chance, yet is considered a relatively conservative estimate of agreement (Hsu & Field, 2003). It can measure the agreement between two raters on binary scales (e.g., how many times two raters both answered “yes” or “no” to a question) or ordinal scales (Cohen, 1968). Fleiss' kappa is an extension of Cohen's kappa and is used to evaluate agreement between multiple raters. The more observations included in the calculation of Cohen’s and Fleiss’ kappa, the more confidently one can produce an estimate of IRA (Landis & Koch, 1977); thus, the number of observations included in the calculation should be as large as possible (Rigby, 2000).

If it is necessary to calculate IRA frequently, and few observations are available to include in a calculation, kappa may not be an adequate metric to gauge agreement. As an alternative, a researcher-derived measure of IRA may be utilized to determine agreement between raters (as was used as in the current study, presented in Appendix A; see Method). In such cases, if IRA is assessed to be too low using a researcher-derived measure, comprehensive intervention implementation training might be introduced to teachers. This is because comprehensive training has been shown to increase agreement
between raters using an indirect measure (e.g., rating scale; Murphy, Garcia, Kerkar, Martin, & Balzer, 1982) and has also been shown to increase agreement between individuals’ self-report ratings and the ratings of “experts” (i.e., trained observers; Sulsky & Balzer; 1988). In general, the amount of training and support implementers require is often underestimated, particularly when it comes to behavior change (for assessment and intervention behaviors; Erchul & Martens, 2010). Thus, thorough, direct training can lead to a variety of improved outcomes in school-based consultation.

**Teacher Training**

Training to promote behavior change can be structured to encourage the acquisition of a variety of skills and be more or less comprehensive. Specifically, training can be indirect (i.e., less comprehensive, didactic, written; LeBel, Kilgus, Briesch, & Chafouleas, 2010) or direct (i.e., more comprehensive, including opportunities to practice that promote appropriate skill use and generalization; Sterling-Turner, Watson, Wildmon, Watkins, & Little, 2001; Schlientz, Riley-Tillman, Briesch, Walcott, & Chafouleas, 2009). Indirect training is typical of consultation in school-based settings in which time is a barrier to direct training (Sterling-Turner et al., 2002). However, when compared to indirect training, there is evidence that direct training can effectively change consultees’ intervention and assessment behaviors (Hiralall & Martens, 1998; Murphy et al., 1982; Sulsky & Balzer; 1988).

**Direct intervention training.** A seminal study investigating the relationship between training and treatment integrity compared the effects of three types of training on participants’ adherence to a multicomponent intervention (Sterling-Turner et al., 2001). Participants received either didactic, modeling, or rehearsal/feedback training on steps of an intervention. For participants in the didactic training condition, researchers
told participants how to implement the intervention and allowed participants to ask questions, but did not demonstrate the intervention nor ask participants to practice the intervention for the purpose of providing constructive feedback. In the modeling training condition, participants were told how to implement the intervention and watched researchers demonstrate how to implement the intervention but were not given the opportunity to ask questions. In the rehearsal/feedback training condition, participants received a verbal explanation of the intervention and were also provided with the opportunity to practice the intervention with researchers. This practice included positive and constructive feedback from researchers about implementation. Results indicated that the use of modeling and rehearsal/feedback training was associated with higher levels of treatment adherence than indirect training. The highest levels of treatment integrity came from participants in the rehearsal/feedback condition, emphasizing the importance of providing implementers with the opportunity to practice and receive feedback before implementing an intervention. The benefit of this procedure may be ensuring that implementers fully understand the intervention procedures, increasing the likelihood that the intervention will be implemented with high treatment integrity (Watson & Robinson, 1996).

Other studies in which explicit, direct intervention training was provided to participants have also produced results that demonstrate improved treatment integrity levels for a variety of study participants: parents (Marchant, Young, & West, 2004), mental health professionals (Perepletchikova & Kazdin, 2005), school-based behavioral consultants (Sheridan, 1992), and teachers (Hiralall & Martens, 1998). As previously stated, results from prior research indicate that teachers often overestimate their adherence to steps of an intervention (Noell et al., 2005, Wickstrom et al., 1998).
However, direct intervention training has been associated with increased adherence to steps of an intervention. Therefore, it may be that when levels of treatment integrity increase after direct intervention training, data collected via direct observation might reflect improved implementation. This in turn may increase IRA with observers and teachers who may be overestimating adherence levels initially.

**Direct assessment training.** Another way IRA might be improved between raters is to offer comprehensive, direct training on how to use a self-report tool. This type of direct training would incorporate explicit practice making ratings and opportunities for feedback from trainers. In one DBR study, participants that engaged in training that incorporated practice making ratings and performance feedback were better able to match ratings made by experts than participants that were trained without opportunities for practice and feedback (Schlientz et al., 2009). If teachers receive direct training on the use of a self-report tool, their intervention implementation may not be affected, but their ratings might change to no longer overestimate adherence. Furthermore, if teachers no longer overestimate adherence, their ratings might more closely approximate those of an observer, therefore improving IRA with an observer’s ratings.

Thus, there is evidence that direct training can affect change by targeting how an intervention is implemented and how to appropriately make ratings when assessing behavior. It has not yet been determined, however, what type of training could most effectively and efficiently increase agreement as exhibited by an observer’s (e.g., a consultant’s) direct observation ratings and teachers’ (e.g., consultees’) self-report ratings of adherence to steps of an intervention. Determining the type of training needed by teachers (e.g., direct training on intervention steps or direct training on the self-report tool) to most effectively and efficiently observe an increase in agreement with an
observer’s ratings may have an impact on recommendations made for data collection in school-based consultation and treatment integrity assessment of evidence-based interventions in educational settings.

**Explicit instruction**

Evidence-based interventions are of increasing importance in school-based consultation, particularly with regard to supporting student outcomes, and multi-tiered frameworks for academic and behavioral support in schools (Kratochwill & Shernoff, 2004). Within multi-tiered frameworks, such as RTI and school-wide positive behavior support (SWPBS), there is an emphasis placed on high-quality instruction and its systematic delivery in the classroom (Chard et al., 2008). High-quality instruction serves as the basis for the integration of evidence-based interventions needed by students who require additional academic and behavioral supports (Sailor, Doolittle, Bradley, & Danielson, 2009). Prior research has indicated that the use of explicit teaching methods is important to the delivery of high-quality instruction (e.g., Rosenshine, 1979). Explicit instruction is teacher-directed, has an academic focus, requires content to be precisely sequenced, and demands that teachers monitor and provide specific corrective feedback to students (Baumann, 1984). In turn, explicit instruction promotes high academic engagement in the classroom (Duffy & Roehler, 1989).

Swanson (1999) identified several additional criteria associated with utilizing explicit instructional strategies, including (a) breaking a task down into smaller steps, (b) breaking instruction into phases, (c) administering probes and student-directed questions, (d) providing visual presentation of instructional material, (e) allowing for independent student practice, and (f) allowing for instruction to be individually-paced. Explicit instruction includes providing several opportunities to respond to material learned and
offering students time to practice using new skills or interact with new material (Horner & Goodman, 2010). The goal of explicit instruction is to transition the responsibility of task completion from instructional personnel to students over time (Joyce, Weil, & Calhoun, 2008).

One method for utilizing explicit instruction is to sequence delivery of instruction in three stages: advance organizer, student enlistment, and post-organizer (Schumaker, 1989). This method is outlined in Appendix B. Within each stage is a series of empirically evaluated instructional practices. In the advance organizer stage, teachers (a) review previous learning by stating what was previously taught and/or asking students to recall material from the last lesson (Mannies, Gridley, Krug, & Grover, 1989; Swanson & Deshler, 2003), (b) define the content of the instructional material by explicitly stating what will be taught and/or the goals of the current lesson (King-Sears & Cummings, 1996), (c) personalize the instructional material to students by discussing the benefits of acquiring the instructional content and suggesting contexts and settings in which the material might be used (Lenz, Alley & Schumaker, 1987), and (d) state lesson procedures or expectations for student outcomes by informing students of the activities in which they will be engaged or the type of performance that is expected (Ellis, Deshler, Lenz, Schumaker, & Clark, 1991).

Students then practice with instructional material in a student enlistment stage. When this stage occurs, teachers (a) prompt involvement by asking students to actively think about the instructional material (MacArthur, Schwartz, Graham, Molloy & Harris, 1996; Englert, 1984); (b) check for understanding of the material taught via pausing during instruction delivery and asking questions that target assessing student comprehension (Fuchs & Fuchs, 1985); (c) expand student responses by correcting and
shaping answers to questions posed (Kline, Schumaker & Deshler, 1991); and (d) engineer students’ success via scaffolding (i.e., breaking down tasks into discrete steps that increase in complexity) and the delivery of positive reinforcement in the form of specific, contextual praise (Palmer, Wehmeyer, Gipson, & Agran, 2004; Rosenshine, 1985).

Finally, teachers initiate a post-organizer stage in which teachers (a) again engage the class in review of instructional material, identifying critical elements of instruction (King-Sears & Cummings, 1996); (b) give direction about what will come next, specifically future material and activities (Faggella-Luby & Graner, 2010; (c) state expectations for future learning and performance, communicating that students are capable of learning and identifying how students will benefit from instruction in the future (Ellis et al., 1991); and (d) personalize the instructional material once more by soliciting student input about generalization of instructional content (e.g., students volunteer when and where they might use the learned content; Lenz et al., 1987).

Examples and non-examples of each step within the three explicit instructional stages described above can be found in Appendix C.

The use of explicit instruction as an instructional intervention has been shown to be effective across content areas (e.g. in science, McCleery & Tindal, 1999; mathematics, Hudson, Miller, & Butler, 2006; social studies, Fontana, Scruggs, & Mastropieri, 2007), as well as when used with emerging readers (e.g., Coyne et al., 2009), English language learners (e.g., Mueller, 2010) and students with learning disabilities (e.g., Witzel, Mercer, & Miller, 2003); however more often it is used in a smaller group setting versus during whole-group instruction. The use of explicit instructional methods, however, is thought to contribute to the delivery of high-quality, effective instruction (Adams & Engelmann,
High-quality, effective instruction has been linked to positive outcomes associated with classroom behavior (Algozzine, Wang, & Violette, 2011) and academic performance (e.g., Caprara, Barbaranelli, Pastorelli, Bandura, & Zimbardo, 2000; Sutherland & Wehby, 2001). Specifically, implementing high-quality, effective instruction in the classroom improves students’ academic engagement (Newcomer, 2009) and decreases disruptive behavior (Nelson, Johnson, & Marchand-Martella, 1996). With vast empirical support, explicit instruction can be used as an intervention to engage in high-quality, effective instructional practice when supports are needed to improve student behavior and academic outcomes in a classroom setting. It is used as the intervention for which ratings are provided in the current study.

The purpose of this study was to evaluate if direct training, specifically direct intervention and direct assessment training, would increase observed IRA between an observer’s and teachers’ self-report ratings of adherence to stages of explicit instruction from an indirect training and screening phase.

**Research Question and Hypothesis**

The study’s primary research questions and hypotheses are as follows.

**Research Questions**

**Research question 1.** Will direct intervention training result in adequate IRA between an observer’s ratings and teachers’ self-report ratings of adherence to stages of the explicit instruction intervention?

**Hypothesis 1.** As previously stated, teachers often engage in overestimation of treatment adherence (Gresham et al., 2000; Noell et al., 2005). Direct intervention training has been associated with increased levels of teachers’ treatment integrity (Sterling-Turner et al., 2001). Increased levels of treatment integrity should minimize
teacher overestimation of intervention adherence, or, in other words, as teachers’ implementation adherence increases, their estimation of adherence should be observed to more closely match their actual behavior. Therefore, it was hypothesized that IRA between an observer’s and teachers’ ratings would improve from an indirect training and screening phase (during which it was hypothesized that IRA would be inadequate) to a direct intervention training phase (during which it was hypothesized that IRA would be adequate).

**Research question 2.** Will *direct assessment training* result in adequate IRA between an observer’s ratings and teachers’ self-report ratings of adherence to stages of the explicit instruction intervention?

**Hypothesis 2.** Teacher self-report ratings of adherence to an intervention have been found to be higher than treatment integrity data collected through direct observation (e.g., Wickstrom et al., 1998). Direct assessment training may help to reduce the observed discrepancy between data collected via these two methods. Simply, after direct assessment training, it is hypothesized that teachers will be less likely to overestimate their adherence when providing self-report ratings. Results from one study indicated that when trainees were offered the opportunity to practice rating using a self-report tool, and were provided with feedback by an expert rater regarding their rating, agreement between trainee and expert rater increased (Schlientz et al., 2009). Therefore, it was hypothesized that IRA between an observer’s and teachers’ ratings would improve from an indirect training and screening phase (during which it was hypothesized that IRA would be inadequate) to a direct assessment training phase (during which it was hypothesized that IRA would be adequate).
Chapter III: Method

Participants

Teacher participants. Three teachers from a midsize urban, Title I school in the Northeast served as participants in the study. The school housed both elementary and middle school grades (i.e., K-8). The first teacher, Teacher A, identified herself as a 29-year-old Asian American female who had been teaching for 4 years total (3 at the current school). She previously taught 3rd grade, but was currently teaching a combined 5th and 6th grade class. During math time, her class was comprised only of 5th grade students from her class and another teacher’s class. The second teacher, Teacher B, did not report her age but identified herself as a multiracial female who had taught three periods of 7th and two periods of 8th grade science for 10 years (6 at the current school). The third teacher, Teacher C, was a 32-year-old Caucasian female who had been teaching a combined 3rd and 4th grade class for 5 years with no teaching experience prior to her current position. During math time, her class was comprised only of 4th grade students from her class and another teacher’s class. All three teachers were certified to teach general education and possessed a Masters degree. Although Teacher A reported having no prior training on implementing instructional or behavioral classroom interventions, Teacher B indicated that she had received some intervention training through professional development activities and Teacher C reported that intervention training had been integrated into her graduate coursework.

Student participants. A total of 64 students were enrolled in participating teachers’ classrooms: 21 in Teacher A’s class (grade 5 math), 22 in Teacher B’s class (grade 7 science) and 21 in Teacher C’s class (grade 4 math). The US Department of Education (2011) reports that half of the 498 students enrolled in the school were
Caucasian (50.2%), a quarter were African American (24.5%), and the remaining students were either Latino (12.5%), Asian or Pacific Islander (7.4%), Native American (<1%) or Multiracial (4.8%). Nearly half of students school-wide were eligible for either free or reduced lunch (45.3%), and many students received special education services (23.7%) and/or spoke a first language other than English (22.9).

**Instrumentation**

**Teacher self-report of adherence.** Teachers rated their adherence to an instructional intervention (i.e., explicit instruction) using the Teacher Daily Instruction Form (see Appendix D). This form had 3 DBR-MISs (described in greater detail below) that served as prompts to teachers to evaluate the stages of explicit instruction: advance organizer, student enlistment and post-organizer stages of instruction (Schumaker, 1989). Each stage within explicit instruction encompassed four steps. Teachers rated their implementation of the advance organizer stage by determining if they (a) reviewed previous learning, (b) defined the content of the lesson, (c) personalized the instructional material to students, and (d) stated lesson procedures or expectations for student outcomes. Teachers rated their implementation of the student enlistment stage by determining if they (a) prompted student involvement, (b) checked for understanding (c) expanded student responses by to questions posed, and (d) engineered students’ success via scaffolding and the delivery of contextual, specific praise. Finally, teachers rated their implementation of the advance organizer stage by determining if they (a) reviewed instructional material, (b) gave direction about upcoming instruction and activities, (c) stated expectations for future learning and performance, and (d) once again personalized the instructional material. Teachers were told to provide one rating for all steps within a stage on a 0-10 point scale and to weigh all steps within the stage equally. This
information was also printed in the directions at the top of the Teacher Daily Instruction Form. It was also explained to teachers that providing a rating of 0 meant that no steps within a stage were implemented whereas giving a rating of 10 meant that all steps within a stage occurred. Ratings were provided immediately after instruction when the primary observer (i.e., student investigator) was present.

**Direct observations of adherence.** The primary observer rated teacher adherence to the intervention using the Teacher Daily Instruction Form. Teachers were typically observed three times a week (although sometimes teachers were observed more or less frequently due to schedule interruptions). The observer looked for all steps of each stage of explicit instruction and weighed all steps within a stage equally. The equal weighting was used as treatment integrity researchers suggest that all steps of an intervention should be weighted equally until more systematic research regarding how to weight steps emerges (Noell, 2008). Therefore, in this study, if a teacher missed a step within a stage, the observers deducted the value of that step from a perfect rating of 10. For example, as there are four steps in each stage of explicit instruction in this study, if a teacher omitted one step (e.g., *review previous learning* in the advance organizer stage) but implemented the other three steps in the stage, the observers deducted 2.5 from a perfect rating of 10. This rating of 7.5, however, was rounded up to 8 to keep ratings consistent with the discrete scale developed for this study and to align with procedures related to the decision tree (see Dependent Variable Definitions and Appendix A).

**Inter-observer agreement.** A secondary observer provided ratings for approximately 30% of the observed intervention sessions for the purpose of calculating inter-observer agreement (IOA) with the primary observer. IOA was determined by dividing the number of agreements (instances in which both the primary and secondary
observer matched ratings or came within one point of each other) by the number of
agreements plus disagreements then multiplying this value by 100 (Berk, 1979; Riley-

Social validity. Finally, social validity was assessed using an adapted version of the
Usage Rating Profile – Intervention (URP-I; Chafouleas, Briesch, Riley-Tillman, &
McCoach, 2009), presented in Appendix E. The URP-I created by Chafouleas and
colleagues (2009) consists of 35 items aimed at reliably assessing the acceptability (α
= .96), understanding (α = .90), feasibility (α = .85), and the perceived need for external or
systems support (α = .84) when implementing an intervention. In this study, the URP-I
was adapted to include multiple sections. The first section prompted teachers to assess
their implementation of explicit instruction and incorporated three items targeting each
construct above: acceptability (items 1, 5 and 9), understanding (items 2, 7, and 10),
feasibility (items 4, 6, and 11), and perceived need for external systems or support (items
4, 8 and 12). The second section prompted teachers to assess receiving direct intervention
training. This section was shorter, incorporating just one item per construct listed above:
acceptability (items 13), understanding (item 14), feasibility (item 15), perceived need for
external or systems support (item 16), as well as an additional item to determine if
teachers would desire to receive direct intervention training again (item 17). A third
section was created to target the social validity of receiving direct assessment training. It
was structured similarly to the second section, incorporating one item per construct listed
above: acceptability (items 18), understanding (item 19), feasibility (item 20), and
perceived need for external or systems support (item 21), as well as an item to determine
teachers’ desire to receive the training again (item 22). (It should be noted, however, that
teachers did not complete this section of the instrument as direct assessment training was
not warranted in the current study. This is described in more detail below). When the measure was adapted, the second and third sections were devised to be shorter than the first section. This is because the URP-I was created to assess intervention usage. Teachers were the recipients of direct training rather than the implementers; however because direct intervention training and direct assessment training were the primary dependent variables of the current study, items targeting these trainings were included in the adapted version of the measure. Teachers completed the adapted measure by responding to all items using a six-point likert scale (1 = strongly disagree and 6 = strongly agree).

Dependent Variable Definitions

Inter-rater agreement. The primary observer and teachers rated implementation of explicit instruction on the Teacher Daily Instruction Form at the end of each lesson in which the primary observer was present. Ratings were compared after each observation to determine if adequate IRA was reached. As described in Chapter II, IRA is typically calculated using percent agreement, Cohen’s kappa (Cohen, 1960; 1968) or Fleiss’ kappa (Fleiss, 1971). In this single-subject, multiple-baseline study, it was necessary to determine IRA on a more frequent basis than is typical in group design studies (i.e., on a formative basis). However, because a repeated measure of agreement was needed that did not require a substantial number of data points for this study, a decision tree (see Appendix A) was developed to guide the determination of IRA between the ratings of the primary observer and teachers across phases of the study. This decision tree allowed for daily comparison of ratings and for a bit of disagreement to still be considered “adequate” IRA. Specifically, the primary observer used the decision tree in the following way. First, the primary observer determined the answer to an initial question: Are any of the
teacher’s ratings (across the three scales) more than 2 points away from the primary observer’s ratings? If so, the teacher was considered to have inadequate IRA for that day. If not, the primary observer determined the answer to a second question: Are the majority of the teacher’s ratings (2/3 of ratings) within 1 point of the primary observer’s ratings? If so, the teacher was determined to have adequate IRA for the day. If not, the teacher was considered to have inadequate IRA for that day.

**Design**

An ABCD multiple baseline across teachers design was used. Using a multiple baseline design required that the intervention be staggered across participants to control for threats to internal validity (e.g., maturation, history; Christ, 2007). All teachers were observed during baseline (Phase A) five times before being provided with indirect training on the instructional intervention, explicit instruction (Phase B). Those teachers who were determined not to be in adequate agreement with the primary observer after five days were screened into the direct intervention training phase using the decision tree logic described above. Direct intervention training (Phase C) was introduced before direct assessment training (Phase D) as its previous evaluation in school-based consultation is consistent with typical consultation practice between a school psychologist and teacher (Sterling-Turner et al., 2001). The three teachers who expressed interest in participating and subsequently took part in baseline and indirect training phases of the study were screened into the indirect training phase due to inadequate IRA. No further recruitment was required, but would have taken place if any teacher had adequate IRA with the primary observer during the indirect training phase.

Teachers screened into the direct training phases were staggered into the direct intervention phase one at a time. The order in which they were staggered was randomly
determined; the student investigator randomized the order of participants using an online research randomizer tool (e.g., www.randomizer.org). The teacher who was randomly assigned to the first baseline received direct training while the other two teachers remained in the indirect training phase. After the first teacher in the direct intervention training phase demonstrated adequate IRA as determined by the decision tree for 3 days, the teacher randomly assigned to the second baseline received direct training; likewise with the teacher randomly assigned to the third baseline. Staggering entry into the direct training phase provided the opportunity for three demonstrations of an effect at three points in time.

**Procedure**

**Recruitment.** Recruitment of school district, administrator, teachers and students was completed as follows.

**District participants.** Nine school districts and 12 individual charter schools were contacted. Of the charter schools, 9 denied the request to conduct research and 3 did not respond to email messages. Of the school districts contacted, 3 approved the request to conduct research and allowed for individual principal contact, 4 denied the request, and 2 did not respond to email messages. For the 3 districts that approved the research, school principals were sent an email describing the purpose and procedures of the study. All principals declined or did not respond to communication except in one district in which two principals agreed to meet with the student investigator to discuss the study.

**School participants.** The student investigator met with both principals individually to discuss the study’s purpose and procedures. After this meeting, the principal decided if she/he consented to staff members’ participation. Both principals
with whom the investigator met approved the research in his/her school. Staff members were then recruited through the principal.

**Teacher participants.** The principals put the student investigator in contact with teachers interested in receiving an instructional intervention due to behavioral and academic concerns in their classrooms. After communicating with interested teachers via email, the student investigator met with five teachers to present details of the study. Teachers had the opportunity to ask questions about the study, volunteer for participation after the meeting commenced, and meet again with the teacher briefly to sign the Teacher Consent Form (see Appendix F). Three teachers returned the Teacher Consent Form.

**Student participants.** Once a teacher agreed to participate, students were provided with a brief verbal explanation about the presence of observers and the study’s occurrence in the classroom. Observers did not interact with students, and observations occurred in such a way that it did not interrupt normal classroom activities. Observers did not know the identity of any students nor collect data on student outcomes as the dependent variable of the study was IRA between the primary observer’s and teachers’ ratings.

**Observer training.** The process by which the primary and secondary observers were trained is outlined below.

**Primary observer training.** A multi-step training occurred for the student investigator to serve as the primary observer and provide expert ratings in the current study. First, the student investigator consulted educational literature to develop an operational definition of all components within each stage of explicit instruction (see Expanded Table of Explicit Instruction Definitions and Sample Dialogue in Appendix C and the Abbreviated Table of Explicit Instruction Definitions in Appendix G). An expert
in explicit instruction and an advisor to the current study then reviewed these operational
definitions for accuracy. With his assistance, the student investigator compiled nine
video clips that illustrated examples of teachers using explicit instruction in the
classroom; three video clips for each stage. The clips depicted a teacher instructing a
mock class and were from university trainings and subsequent research studies. The
student investigator then edited the selected clips to be about 1-3 minutes long. Some
clips showed all steps within an explicit instruction stage and other clips depicted only
some steps within a stage.

Then, five experts in explicit instruction were contacted for assistance with
training the student investigator to become the primary observer in May 2011. Each
expert was sent a materials package. In the package was an instruction sheet that
prompted experts to review the three stages of explicit instruction (i.e., advance
organizer, student enlistment, and post-organizer stage) presented in a table included in
the packet (Expanded Table of Definitions and Sample Dialogue; see Appendix C). It
was acknowledged that the definitions provided in the table might not necessarily align
with the experts’ own definitions of explicit instruction; however, experts were asked to
consider only the definitions provided as they represented behaviors targeted for data
collection in the current study. Experts were also given a quick reference of the
operational definitions in a table included in the packet (Abbreviated Table of Explicit
Instruction Definitions; see Appendix G).

Once experts reviewed instructions, operational definitions and examples, they
were asked to watch the nine video clips included on a DVD in the materials package.
Each video clip depicted some or all steps within a stage of explicit instruction (i.e.,
advance organizer, student enlistment or post-organizer stage). At the end of each video
clip, experts were prompted to provide a single rating (from 0-10) for the stage depicted in the clip on a DBR-MIS identical to those used on the Teacher Daily Instruction Form. Experts were told to weigh all steps within a stage equally when making a rating. Specifically, they were instructed that if a teacher did not implement a step within a stage during a video clip to deduct 2.5 points from a perfect rating of 10. The task was estimated to take experts approximately 15-20 minutes to complete. Experts were asked to return all materials to the student investigator (i.e., primary observer) by the beginning of July 2011.

Experts were made aware that ratings provided for each clip would be compared to the ratings of the primary observer to assist in developing an expert consensus for use in the observer’s dissertation. It was further explained that expert consensus was needed so that the primary observer might confidently provide ratings when observing teachers’ instruction as part of data collection. Of the five experts contacted, four participated in the expert consensus procedure. A list of these experts and their credentials can be found in Appendix H.

The median value ascribed to each video clip was used as the expert consensus for that clip. For the student investigator to serve as the primary observer, the student investigator independently watched the 9 video clips, provided her own ratings for each clip, and compared her ratings to that of the expert consensus. When disagreements arose between a rating given by the student investigator and the experts, the student investigator re-watched the clip and provided a new rating until a rating that matched the experts was provided. The ratings from experts were relevant to training not only the primary observer, but also the secondary observer.
**Training the secondary observer.** The student investigator met with the secondary observer, a graduate student in school psychology from a large, public university in the Northeast, for a one-hour training in January 2012, prior to the start of data collection. First, the primary observer described DBR to the secondary observer using a training presentation developed by Chafouleas, Riley-Tillman, and Jaffery (2011). Defining features of DBR were explained as well as its use in educational interventions (e.g., home-school notes). The student investigator then explained how to rate on a DBR-MIS and introduced the secondary observer to the Teacher Daily Instruction Form (Appendix D). The steps of each explicit instruction stage (advance organizer, student enlistment, post-organizer) were pointed out on each of the corresponding scales before the student investigator described each step using the Expanded Table of Explicit Instruction Definitions and Sample Dialogue document (Appendix F). As previously described, this document presented an operational definition and sample dialogue for each step. Then, the primary observer and secondary observer reviewed the Abbreviated Table of Explicit Instruction Definitions (see Appendix G) to verify that the secondary observer understood the definitions of each explicit instruction stage. The student investigator explained that this document could also be used as a “Quick Guide” during observations to remind the secondary observer of operational definitions for each step.

After introducing DBR, the Teacher Daily Instruction Form, and operational definitions for steps of explicit instruction, the primary observer and secondary observer watched a video clip of each stage of explicit instruction twice. The video clips used for this training were the same as those that were utilized for the expert consensus procedure to train the primary observer. During this training, each clip was initially viewed in its
entirety with no interruptions. Then, the primary observer played it again and stopped the clip after each step in a stage was demonstrated to identify that step to the secondary observer. After all clips were viewed, the secondary observer indicated she was ready to practice making ratings and comparing these ratings to those that were devised by experts. The secondary observer watched a clip then made a rating. This rating was compared to the experts’ median rating. The secondary observer was in agreement with the experts’ rating for the first video (a clip of an advance organizer stage), but not the second (a clip of a student enlistment stage). The definitions of steps in that stage were again reviewed with her until the secondary observer was confident in her ability to make ratings. The secondary observer watched the clip again as well as four more clips (one depicting an advance organizer stage, one depicting a student enlistment stage, and two depicting post-organizer stages) and was in 100% agreement with the ratings provided by experts for all 6 clips viewed.

**Baseline.** All participants began in the baseline phase for 5 days (Kratochwill et al., 2010). The primary observer typically watched classroom instruction during the period selected by the participant three times per week (although due to scheduling interruptions and rescheduling, occasionally observations occurred two or four times a week) and rated teachers’ adherence to stages of the instructional intervention using the Teacher Daily Instruction Form. Teachers did not self-report their behavior until the indirect training and screening phase (i.e., Phase B). Therefore, IRA was not calculated during baseline. The purpose of this phase was to determine teachers’ preliminary rates of implementation of explicit instruction. After baseline, all teachers’ movement into the indirect training phase occurred simultaneously as indirect training was not the independent variable.
**Indirect training and screening.** After 5 days of baseline, all teachers met with the primary observer individually to receive indirect training on explicit instruction. In indirect training, teachers are typically told about an intervention via didactic instruction and/or provided with written intervention materials only (Sterling-Turner et al., 2001). In this study, indirect training was didactic and involved the primary observer offering teachers a concise overview of explicit instruction procedures as well as a brief description of each step within the three stages of the intervention. Then, teachers learned how to use the Teacher Daily Instruction Form by receiving a brief training on completing a DBR form (modeled after Chafouleas, Riley-Tillman, & Sugai, 2007). Teachers were told that after each lesson during which the primary observer was present, they would be asked to rate the completeness with which the three stages of the explicit instruction intervention were implemented as planned on the Teacher Daily Instruction Form. This type of training was evaluated prior to direct intervention training because, as previously stated, it is typical of consultation in school-based settings in which time is a barrier to direct training (Sterling-Turner et al., 2002).

Indirect training did not include components of direct training such as modeling, role playing, rehearsal, as well as positive or corrective feedback (Perepletchikova, & Kazdin, 2005; Sterling-Turner et al., 2001), which were provided in the direct training phase (Phase C). Appendix I contains the script that the student investigator used in indirect training with teachers. Once teachers received the indirect training, they were asked to start using explicit instruction daily in their classroom. They were told that after a few weeks, they may meet with primary observer again to talk more about the intervention.
After indirect training, the teachers chose one period after which they would complete Teacher Daily Instruction Form. Teachers C and A taught multiple subjects, but targeted math for implementation of explicit instruction and ratings (Teacher C only taught science, so picked one of her five classes to target). During this phase, approximately three times per week, the primary observer watched teachers’ implementation of explicit instruction and coded adherence using the three DBR-MISs. After each observation session, the primary observer compared her ratings to the teacher’s and utilized the decision tree (see Appendix A) to determine if participants would continue in the indirect training phase or would proceed to direct training. Before the decision tree was used, each potential participant remained in the indirect training and screening phase for 5 days as 5 data points is the minimum number required per phase by What Works Clearinghouse (Kratochwill et al., 2010). After the fifth data point was collected, the primary observer determined the level of IRA between the observer’s and the teacher’s ratings on the three scales for the five days.

Teachers had to demonstrate 3 consecutive observations for which inadequate IRA occurred, as defined by the decision tree (see Appendix A), to be eligible for direct training. Three eligible teachers were needed before moving to direct training to meet the minimum requirements of a multiple baseline design. Once three teachers were eligible to receive direct training, and at least 5 data points had been collected in the indirect training and screening phase, teachers who are eligible to receive direct training were staggered into a direct training phase. As mentioned above, the three teachers that began the study were all eligible to receive direct intervention training based on ratings provided in the indirect training and screening phase (see Figures 1, 2, and 3).
**Direct intervention training.** Teachers entering the direct intervention training phase (Phase C) received a comprehensive training, guided by a script (see Appendix J) and a video presentation (based on slides presented in Appendix K). A script and a video presentation were used so that training was standardized across participants. The video presentation reviewed the stages of explicit instruction in detail and provided a model of explicit instruction for teachers. The script guided the student investigator to ask teachers to participate in a role-play of explicit instruction procedures. The student investigator subsequently provided positive and/or corrective feedback to the teacher. Modeling, role-playing, and providing corrective feedback are procedures consistent with direct training as described in Sterling-Turner et al. (2002).

Once a teacher completed the direct intervention training, she was asked to implement explicit instruction in the classroom and was told that she might meet with the student investigator once more to discuss how to make ratings on the Teacher Daily Instruction Form. The primary observer continued collecting adherence data approximately three times a week, as was done during the indirect training phase (Phase B). After each observation session, teachers and the primary observer completed the Teacher Daily Instruction Form. The primary observer then used the decision tree to determine IRA between ratings provided by both the primary observer and the teacher.

Teachers had to demonstrate 3 consecutive days of inadequate IRA, as defined by the decision tree (see Appendix A), to be eligible for direct assessment training. If teachers did not demonstrate 3 consecutive observations for which inadequate IRA was calculated during the direct intervention training phase, they remained in the direct intervention phase (and therefore did not receive direct assessment training) for a minimum of six observations.
**Direct assessment training.** The investigator prepared a script (see Appendix L) and video presentation (based on slides presented in Appendix M) for direct assessment training. The video presentation focused on reviewing DBR as well as procedures involved in rating teacher behavior (LeBel et al., 2010) using the Teacher Daily Instruction Form in greater detail than was provided in indirect training.

Specifically, the training was designed to begin with teachers (a) receiving an overview of the use of DBR, (b) being provided with operationally defined target teacher behaviors in explicit instruction, and (c) being shown examples and non-examples of ratings that correspond to clips of explicit instruction. Then, teachers were to view a series of short video clips of explicit instruction. At the end of each clip, teachers were to rate stages of explicit instruction on the provided Teacher Daily Instruction Form. Teachers’ ratings would then be compared to ratings predetermined by the experts in explicit instruction solicited for expert consensus training. Finally, the investigator intended to play each clip a second time so that the teacher could see the video clip again in close proximity to learning the value ascribed to each video by experts’ rating. These training procedures are consistent with those used in Schlientz et al. (2009). As presented in the results, the direct assessment training did not occur due to participants’ responsiveness to the direct intervention training. As noted above, direct intervention training was introduced first as its previous evaluation in school-based consultation (Sterling-Turner et al., 2001) is consistent with typical consultation practice between a school psychologist and teacher.

**Deviations to study procedures.** To prepare for any deviations to study procedures that may have impacted data collection (e.g., school cancellation, teacher absence), pre-determined solutions to potential scheduling disruptions were developed
(see Appendix N) and reviewed with the secondary observer and teachers who agreed to participate in the study.

Data Analysis

To analyze results, agreement between the primary observer’s and teachers’ ratings for the three stages of explicit instruction was analyzed via visual analysis, the calculation of effect sizes, and a statistical measure of IRA. Specifically, visual analysis of the trend, level, and stability of data paths was conducted to determine the presence of a functional relationship between providing direct intervention training and IRA. As a multiple baseline design was used, three replications of effect were needed to determine if a functional relationship was evident (e.g., Horner et al., 2005).

Additionally, calculation of three non-parametric statistics to determine the effect size (i.e., to quantify the magnitude of the difference) of IRA across training phases was used: percentage of non-overlapping data (PND), standard mean difference (SMD) and improvement rate difference (IRD). This was accomplished using the difference between ratings furnished by the primary observer and teachers. Three estimates were used in interpretation of results to account for the strengths and limitations of each statistic.

For PND calculations, the number of non-overlapping data points between the indirect training phase and direct intervention training phase was divided by the total number of data points in the direct intervention training phase to produce a percentage of non-overlapping data. This method was used due to previous support for its consistency with visual analysis data despite limitations related to lack of consideration for trend, outliers and magnitude of effect (Wolery, Busick, Reichow, & Barton, 2010). When interpreting PND, it is suggested that PND, >.90 is very effective, .70 - .90 is effective, .50 - .70 is questionable and <.50 is ineffective (Scruggs & Mastropieri, 1998).
For SMD calculations, the mean of the direct intervention training phase was subtracted from the mean of the indirect training phase and divided by the standard deviation of the indirect training phase to produce an effect size estimate. For IRD, an “improvement rate” is calculated for both indirect training and direct intervention training phases. An improvement rate for indirect training is calculated by dividing the number of data points that meet or “exceed” (in this case, fall below) any direct intervention training data point by the total number of points in the indirect training phase. An intervention improvement rate is calculated by dividing the number of data points that “exceed” all indirect training data points by the total number of points in the direct intervention training phase. Then the direct intervention training improvement rate is subtracted from the indirect training improvement rate to determine the percentage of overlapping data points across phases (i.e., IRD estimate; Parker, Vannest, & Brown 2009). When interpreting IRD, it is suggested that a value > .50 exhibits a small effect, .50 - .70 exhibits a moderate effect and above .70 is a large effect (Parker et al., 2009). Calculation of SMD and IRD accounts for magnitude of effect, and may account for outliers, but trend remains unconsidered in the estimates (Busk & Serlin, 1992; Parker et al., 2009).

Finally, the calculation of Cohen’s kappa (Cohen, 1960), a fairly conservative statistical index of rater consensus that may be lower than other agreement calculations (e.g. IRA using a decision tree), was calculated to provide a quantitative comparison of IRA across training phases. To generate the kappa statistics, three pairs of ratings from each observation were compared: one pair for the advance organizer, one pair for the student enlistment, and one pair for the post-organizer stage. Subtracting the probability of chance agreement (Pr_c) by the observed agreement (Pr_a) among raters, then dividing that value by 1 minus the probability of chance agreement (Pr_c) was used to calculate
kappa (i.e., \( k = \frac{Pr_a - Pr_e}{1 - Pr_e} \)). To interpret kappa, Umesh, Peterson, and Sauber (1989) suggested taking the ratio of kappa to kappa maximum, the maximum value for observed agreement, \( k_{\text{max}} = \frac{P_{\text{max}} - P_{\text{exp}}}{1 - P_{\text{exp}}} \), to determine the observed proportional agreement between raters (i.e., \( k / k_{\text{max}} \)).
Chapter IV: Results

Figures 1-6 present data related to IRA between raters. The two data paths in Figures 1, 2 and 3 depict the primary observer’s and teachers’ ratings for explicit instruction in the advance organizer, student enlistment, and post-organizer stages, respectively. Figures 4, 5 and 6 depict the magnitude (i.e. number of points) of the difference between teacher and primary observer ratings across phases. Effect size and Cohen’s kappa calculations are presented in Tables 1 and 2 respectively. These calculations were generated to support the interpretation of results from visual analyses. Furthermore, teacher’s average adherence to stages of explicit instruction across phases is presented in Table 3 and described below. This is followed by results from inter-observer agreement calculations and the administration of the URP-I for social validity data.

IRA

Overall, during the indirect training phase (Phase B), the difference between ratings provided by the primary observer and teachers ranged from 0-7 and averaged 2.76 points ($SD = .94$). In the direct intervention training phase (Phase C), the difference between ratings provided by the primary observer and teachers again ranged from 0-7, but averaged .62 points ($SD = .49$). This builds support for the observation that IRA between the primary observer and teachers improved after direct training was provided to teachers. For Teachers C and A, after direct intervention training, the primary observer provided higher ratings for intervention implementation across stages of explicit instruction. This increased agreement between ratings provided by the primary observer and teachers. After Teacher B received direct intervention training, her ratings decreased on the advance organizer, student enlistment and post-organizer scales. This increased
her agreement with the primary observer whose ratings did not change from the indirect training phase.

These changes in agreement across phases are perhaps more evident in Figures 4, 5, and 6 which depict the difference between ratings provided by the primary observer and teachers using a single data path. These data were used to calculate effect size estimates (see Table 1), which showed varied estimates of a treatment effect by explicit instruction stage. Finally, the calculations of Cohen’s kappa, kappa maximum, and the ratio of kappa to kappa maximum presented in Table 2 illustrate that overall, the primary observer and teachers had a higher degree of consensus after direct intervention training.

**Teacher C.** Teacher C received indirect training on 2/7 and remained in the indirect training phase for 5 observations (2/7, 2/8, 2/13, 2/14, and 2/15; the teacher canceled an observation on 2/9 to give students a unit assessment). During the indirect training phase, Teacher C’s ratings were consistently high and stable for the advance organizer stage (see Figure 1; \( M = 8, SD = 0 \)), high and slightly variable for the student enlistment phase (see Figure 2; \( M = 7.6, SD = .55 \)) and high but more variable for the post-organizer stage (see Figure 3; \( M = 7, SD = 1.22 \)). The primary observer’s ratings showed similar variability across instructional stages, but tended to be lower for the advance organizer (\( M = 5; SD = 0 \)) and post-organizer stages (\( M = 4.6; SD = 2.88 \)) and higher for the student enlistment stage (\( M = 9.2; SD = 1.10 \)). Using the decision tree, it was determined that the primary observer and Teacher C were in adequate agreement for 0 of 5 observations in the indirect training phase.

Teacher C received direct intervention training on 2/17 and remained in the direct intervention training phase for 10 observations (2/28, 3/1, 3/5, 3/7, 3/9, 3/13, 3/14, 3/20, 3/27 and 4/3; 2/16-2/28 was spring break and observations were canceled on 2/29 due to
teacher illness, 3/6 due to a half day schedule, 3/29 due to state testing, 3/15, 3/21, 3/22, and 3/28 by the primary observer to make up observations with another teacher, and 3/30 due to the primary observer attending a conference). In the direct training phase, Teacher C’s ratings were consistently high and slightly variable for the advance organizer ($M = 9.2; SD = .63$) and student enlistment stages ($M = 8.8; SD = .63$) and moderately high but more variable in the post-organizer stage ($M = 7.2; SD = 3.88$). Although ratings provided by the primary observer showed similar variability as the ratings furnished by Teacher C across stages, the primary observer’s ratings were consistently higher for the advance organizer ($M = 10; SD = 0$), student enlistment ($M = 8.8; SD = 1.03$) and post-organizer ($M = 9.6; SD = .84$) stages. No clear trend was detected for ratings provided by both the primary observer and Teacher C in either the indirect or direct intervention training phase. Teacher C and the primary observer were determined to be in adequate agreement for 9 of the 10 observations in the direct intervention training phase. The one exception was on 3/27 when Teacher C provided a rating of 8 on the advance organizer and post-organizer scales and the primary observer provided a score of 10 to both scales.

Upon examining the data paths in Figures 1, 2 and 3, visual analysis suggests that agreement between Teacher C and the primary observer increased in the direct intervention training phase from the indirect training phase for all three stages of explicit instruction. Some of the effect size calculations presented in Table 1 support these results. PND calculations indicated that, for agreement between raters, a moderate or large effect size was found for the advanced organizer stage (PND = 1.00), student enlistment (PND = 1.00) and post-organizer (PND = .70) stages. IRD calculations also indicated that a large effect size was found for the advance organizer stage (IRD = 1.00), but that small effect sizes were produced for the student enlistment (IRD = -.40) and post-
organizer (IRD = -0.10) stages. SMD calculations were undefined for the student enlistment stage (as the SD for the baseline phase was 0, thus dividing the difference in means by the SD of the baseline resulted in a quotient that was undefined), but moderate to large for the student enlistment (SMD = -1.06) and post-organizer (SMD = -1.84) stages.

Calculations of Cohen’s kappa presented in Table 2 indicate that consensus between ratings provided by Teacher C and the primary observer was low in the indirect training phase ($k = -0.04$). Given the maximum possible agreement ($k_{\text{max}} = 0.10$), the ratio of kappa to the maximum possible kappa demonstrated fair agreement ($k/k_{\text{max}} = -0.41$). That is, the observed value of kappa was 41% as large as it possibly could be, given the circumstances (e.g., number of observations). In the direct intervention training phase, the ratio of kappa to the maximum possible kappa demonstrated high agreement ($k/k_{\text{max}} = 0.86$). In this phase, the value of kappa was 86% as large as it possibly could be, given the circumstances.

**Teacher B.** Teacher B received indirect training on 1/31 and remained in the indirect training phase for 8 observations (2/6, 2/7, 2/10, 2/13, 2/14, 2/17, 2/27 and 2/28; the teacher canceled an observation on 3/5 due to an out-of-school training). During the indirect training phase, Teacher C’s ratings were moderate to high and somewhat variable for the advance organizer stage ($M = 9, SD = 1.20$), moderate to high, somewhat variable, and with an increasing trend for the student enlistment phase ($M = 8.63, SD = 1.41$) and high and fairly stable for the post-organizer stage ($M = 9.5, SD = 0.93$). No clear trend was noticed for the advance organizer and post-organizer stages. The observer’s ratings were moderate and somewhat variable for the advance organizer ($M = 5.25, SD = 1.91$) and moderate to high and somewhat variable for the student enlistment stage ($M = 6.88, SD = 1.73$).
Although no clear trend was detected in ratings made in either stage. For the post-organizer stage, ratings were moderate to high and stable with an increasing trend ($M = 6.13; SD = 1.55$). Use of the decision tree indicated that the primary observer and Teacher B were not in adequate agreement for any day of the indirect training phase.

Teacher B received direct intervention training on 3/6 and remained in the direct intervention training phase for 10 observations (3/6, 3/8, 3/5, 3/12, 3/15, 3/16, 3/19, 3/22, 3/23, 3/26 and 4/3; observations were canceled on 3/9 due to an in-class assessment and on 3/13 due to a field trip). In the direct training phase, Teacher B’s ratings were moderate to high, variable, and with no clear trend for the advance organizer ($M = 7.4, SD = 1.58$), student enlistment ($M = 8.8, SD = .63$) and post-organizer ($M = 6.6, SD = 2.88$) stages. Teacher B provided a rating of 1 and 0 for the student enlistment and post-organizer stages respectively on 3/12 as a guest speaker presented during the second half of the class period. Overall, Teacher B’s ratings of adherence to the three stages were lower than in the indirect training phase. Compared to Teacher B, ratings provided by the primary observer were slightly higher, showed similar variability and also demonstrated no clear trend for the advance organizer ($M = 7.3, SD = 2.26$), student enlistment ($M = 7.3, SD = 2.91$) and post-organizer ($M = 5.8, SD = 2.74$) stages. Using the decision tree, the primary observer and Teacher B were in adequate agreement for 8 of 10 days of the direct training phase.

As shown in Figures 1, 2 and 3, agreement between Teacher B and the primary observer increased in the direct training phase from the indirect training phase for all three stages of explicit instruction. The effect size calculations support these results. PND calculations indicated that, for agreement between raters, there was a large effect size in the advanced organizer (PND = 1.00) and post-organizer (PND = .80) stages, and a
moderate effect size in the student enlistment (PND = .60) stage. SMD calculations indicated a large effect size for agreement between raters in the advance organizer (SMD = -.88), student enlistment (SMD = -1.11), and post-organizer stage (SMD = -1.29). Finally, IRD calculations indicated a small effect size for the advance organizer (IRD = -.75), student enlistment (IRD = -.15) and post-organizer (IRD = -.20) stages.

Calculations of Cohen’s kappa presented in Table 2 indicate that consensus between ratings provided by Teacher B and the primary observer was low in the indirect training phase ($k/k_{max} = -.30$). That is, the observed value of kappa was 30% as large as it possibly could be, given the circumstances. In the direct intervention training phase, the ratio of kappa to the maximum possible kappa demonstrated high agreement ($k/k_{max} = .74$). In this phase, the value of kappa was 74% as large as it possibly could be, given the circumstances.

**Teacher A.** Teacher A received indirect training on 2/1 and remained in the indirect training phase for 13 observations (2/3, 2/6, 2/7, 2/14, 2/16, 2/17, 2/27, 2/28, 3/2, 3/6, 3/9, 3/12 and 3/16; the teacher canceled an observation on 2/10 due to illness). During the indirect training phase, Teacher A’s ratings were typically high and somewhat variable (with no clear trend) for the advance organizer stage ($M = 8.46, SD = 1.05$), student enlistment ($M = 8.31, SD = 1.25$) and post-organizer ($M = 8.38, SD = .87$) stages. The primary observer’s ratings were moderate to low and variable for the advance organizer ($M = 4.69; SD = 1.75$) and post-organizer ($M = 5.08; SD = 1.01$) stages, but high and less variable for the student enlistment ($M = 9.23; SD = 1.01$) stage. Use of the decision tree indicated that the primary observer and Teacher A were in adequate agreement for 0 of 13 observations in the indirect training phase.
Teacher A received direct intervention training on 3/16 and remained in the direct intervention training phase for 6 observations (3/21, 3/27, 3/28, 4/2, 4/4, 4/5; observations were canceled on 3/19, 3/20, 3/22, 3/23 and 3/29 due to state testing, 3/30 due to a conference and on 4/3 due to teacher illness). In the direct training phase, Teacher A’s ratings were high with little to no variability and no trend for the advance organizer \((M = 10, SD = 0)\), student enlistment \((M = 9.67, SD = .82)\) and post-organizer \((M = 8, SD = 4)\) stages. Teacher A provided a rating of 0 for the post-organizer stage on 4/4 as she ran out of time to conduct the steps of the post-organizer stage in her lesson. Overall, Teacher A’s ratings for the three stages were higher than in the indirect training phase. Compared to Teacher A, ratings provided by the primary observer were nearly identical for all three stages: the advance organizer \((M = 10, SD = 0)\), student enlistment \((M = 9.83, SD = .41)\) and post-organizer \((M = 8.33, SD = 4.08)\) stages. Using the decision tree, the primary observer and Teacher A were in adequate agreement for all 6 days of the direct training phase.

Like Teachers C and B, results of visual analysis suggests that agreement between Teacher A and the primary observer increased in the direct training phase from the indirect training phase for all three stages of explicit instruction. The effect size calculations show mixed these results. PND calculations indicated that, for agreement between raters, there was not an effect across the advance organizer (PND = .00) student enlistment (PND = .00) and post-organizer (PND =.00) stages. IRD calculations indicated a small or moderate effect across the advance organizer (IRD = -.08) student enlistment (IRD = -.54) and post-organizer (IRD = .53) stages. However, SMD calculations produced moderate to large effect sizes for agreement between raters in the advanced
organizer (SMD = -.97), student enlistment (SMD = -1.26), and post-organizer (SMD = -1.75) stages.

Calculations of Cohen’s kappa presented in Table 2 indicate that consensus between ratings provided by Teacher A and the primary observer was low in the indirect training phase \( (k/k_{max} = 0.03) \). That is, the observed value of kappa was 3% as large as it possibly could be, given the circumstances. In the direct intervention training phase, the ratio of kappa to the maximum possible kappa demonstrated near perfect agreement \( (k/k_{max} = 1.0) \). In this phase, the value of kappa was about 100% as large as it possibly could be, given the circumstances.

**Adherence**

Overall, teacher’s adherence to steps within the three stages of explicit instruction increased from baseline in both the indirect and direct intervention training phases. As depicted in Table 3, the average rating provided by the primary observer for teachers’ implementation of steps of the advance organizer stage increased across phases: 4.6 during baseline, 5 during the indirect training phase, and 9.1 during the direct intervention training phase. This pattern was similar for ratings provided by the primary observer for teachers’ implementation of steps of the student enlistment stage across phases: 7.53 during baseline, 8.43 during the indirect training phase, and 8.73 during the direct intervention training phase. An increase also occurred for ratings provided by the primary observer for teachers’ implementation of steps of the post-organizer stage across phases: 3.26 during baseline, 5.27 during the indirect training phase, and 7.73 during the direct intervention training phase.

**Inter-observer Agreement**
The secondary observer accompanied the primary observer to 20 of 67 total observations (29.9%). Overall, inter-observer agreement between the primary observer and secondary observer ranged from 75-100% and averaged 92.5% (SD = 6.6). Inter-observer agreement data by teacher and observation date are presented in Table 4.

Social Validity

Overall, results from the URP-I, presented in Table 5, indicate that teachers found explicit instruction and direct intervention training to be favorable interventions. Specifically, all three teachers provided high ratings for items related to the acceptability (M = 5.67), feasibility (M = 5.89), and their understanding (5.57) of explicit instruction. Ratings were typically low for teachers’ perceived need for external or systems support to implement explicit instruction (M = 2.22). For direct intervention training, ratings for items measuring acceptability (M = 5.67), feasibility (M = 6), and understanding (M = 5.67) were high and perceived need for supports (1.33) were low, but Teacher B indicated that she would most likely not desire to receive direct intervention training again whereas Teachers A and C indicated that they would. Anecdotally, Teacher A said that she, “thought [explicit instruction] was very easy to imbed in what she was already doing” and that it was “like spice… it added flavor to [her] instruction”. Teacher C said that she “really thought it helped her practice and that she “will keep using [explicit instruction] and have talked to other teachers about it, too”.

46
Chapter V: Discussion

Federal legislation and educational policy have led school personnel to adopt multi-tiered service delivery frameworks to support students’ academic and behavioral outcomes. When implementing plans and procedures within such a framework, treatment integrity data are critical pieces of information to consider when evaluating a students’ responsiveness to an intervention. Collecting treatment integrity data can be done via direct or indirect methods, but although direct methods have been shown in research to be more reliable, they are resource-intensive and perhaps impractical for consistent use in school settings. Indirect methods, such as teacher self-report, require less time for school psychologists to coordinate, but may result in teachers’ overestimating their adherence to an intervention and not align with the data that a direct observer might collect (i.e., result in low IRA between the two raters). In prior studies, it has been unclear as to why teachers might overestimate their adherence; however it may be due to a lack of training (Erchul & Martens, 2010) specifically to implement the intervention or use the self-report tool as results from prior research have lent support to the possibility that comprehensive training can increase adherence to steps of an intervention (Hiralall & Martens, 1998; Sterling-Turner et al., 2001) and IRA between raters (Schlientz et al., 2009). Therefore, the purpose of this study was to determine if direct training would increase IRA between an observer’s and teachers’ self-report ratings of adherence to stages of explicit instruction from an indirect training phase.

Results suggest that the consistent overestimation of treatment adherence provided by teacher ratings in the indirect training phase in this study is similar to that which has been observed in previous research (e.g., Wickstrom et al., 1998). Although in this study, teachers frequently rated their implementation 2-7 points higher than the
primary observer, sometimes a teacher (e.g., Teacher A implementing the student enlistment stage) rated her adherence lower than the primary observer, supporting results from prior research that indicates teachers often but not always overestimate adherence (e.g., Sanetti & Kratochwill, 2009b). In such cases, however, because ratings provided by the primary observer and teachers were consistently dissimilar, IRA was determined to be inadequate for most observations across teachers during the indirect training phase.

After direct intervention training, increased levels of adherence to steps of the intervention occurred. This is also consistent with results from previous studies (Sterling-Turner et al., 2001). Results of this study extend the line of research on direct training in that they indicate there may be a relationship between direct intervention training and an increase in IRA between the primary observer’s and teachers’ self-report ratings of adherence to stages of explicit instruction, as hypothesized in Chapter II. For two teacher participants, Teachers C and A, results support the hypothesis that as teachers’ implementation levels increased post-direct intervention training, their self-report ratings also shifted from overestimating their adherence in the indirect training phase to more closely matching their actual behavior in the direct intervention training phase.

For Teacher B, adequate IRA was achieved for all days in the direct intervention training phase. That is, once provided with direct intervention training, Teacher B’s implementation of the intervention steps improved a bit (like Teachers C and A, but to a lesser extent), but also appeared to impact on her ratings. She lowered her adherence ratings to more closely match those of the primary observer’s after the direct training. This provides important initial evidence that there may not only be a relationship between direct intervention training and (a) increases in intervention adherence, (b) IRA between raters, but also (c) a reduction in overestimation when teachers rate their adherence to
steps of an intervention. This last finding may be an important relationship to investigate in future research as systematic overestimation (or fixed bias; Del Boca & Noll, 2000) has limited the use of teacher self-report in previous treatment integrity studies, despite it being commonly used to collect treatment integrity data in the field (Cochrane & Laux, 2008). Further research examining the relationship between direct intervention training and teacher self-report ratings may begin to bridge the gap between research and practice.

To gauge the relationship between direct intervention training and IRA in the current study, three effect size calculations (PND, SMD, and IRD) were utilized, but results were mixed. For PND, moderate to large effect sizes were noted, supporting results from visual analysis (a strength attributed to the statistic in previous research; Wolery et al., 2010), as well as kappa calculations presented in Table 2. As the statistic does not account for outlier data points, a limitation mentioned in research by Wolery and colleagues (2010), it appears that no effect was achieved in Teacher A’s classroom. This result, however, was not supported by the calculation of SMD.

Moderate to large effect sizes were noted upon calculating SMD because outliers do not affect the calculation of the statistic in the way that it can when calculating IRD (Busk & Serlin, 1992; Parker et al., 2009). As such, effect sizes as determined by IRD were comparatively small. Yet, visual analysis supports the larger effect sizes found by calculating PND and SMD. Overall, these results demonstrate the need for further research around determining or developing an appropriate effect size for examining the relationship between IRA and an intervention such as direct intervention training.

Explicit instruction was chosen as the intervention for which teachers received training due to its generalizability across subjects and settings. That is, explicit instruction reflects best practice in instructional pedagogy and is appropriate for use in any
classroom. This was important in the current study as recruitment of teachers across grade levels and subjects was necessary. Future research might evaluate the efficacy of the components of explicit instruction utilized in this study across other settings.

As described in the Method, participants received direct intervention training on explicit instruction first due to its previous evaluation in school-based consultation literature and because it is more typical of consultation practice between a school psychologist and teacher (Sterling-Turner et al., 2001). Teachers were to receive direct assessment training if direct intervention training failed to produce adequate IRA for three consecutive days. However, due to the responsiveness of teachers to direct intervention training, direct assessment training was not warranted and therefore not implemented. Future research might assess if direct assessment training produces an effect similar to that which was achieved after direct intervention training with the three teachers in the current study. It is possible that the two trainings were not a bifurcated as was intended in the study. That is, teachers may have become better raters after direct intervention training, even though that was not the intention of direct intervention training. To further investigate this possibility, more research is needed. If both trainings are offered in future studies, information regarding outcomes can be analyzed with social validity data to help determine the type of training that works best to increase IRA, but also, account for what teachers prefer. In this study, teachers found both explicit instruction and direct intervention training to be acceptable and feasible, understood both concepts, and rated their perceived need for support to be low. Future research might assess if these ratings are replicated in similar studies of support and training provided through a school consultation framework.
In consultation, as treatment integrity data are often not collected due to a lack of time (Cochrane & Laux, 2008), and other resources may be limited, maximizing efficiency is paramount. Typically, treatment integrity data may be collected via direct observation. This method produces more reliable data than one alternative, teacher self-report, but is time-intensive (e.g., observing the intervention used in the current study for just one teacher over six weeks required 360 minutes of time). Collecting teacher self-report data is less resource-intensive than direct measures (Hersen, 2004; Riley-Tillman et al., 2008), may be seen as less evaluative than direct observation (Biggs et al., 2008), and might provide teachers with the opportunity to gain skills in data collection, but may also produce data that are less reliable than direct observation (Wickstrom et al., 1998).

The time required for offering teachers direct intervention training in this study was 45 minutes. After that time, teachers’ implementation adherence increased and IRA with the primary observer improved to be adequate across participants. With the bit of time invested in comprehensive intervention training, individually or perhaps in small group settings (the latter being a worthy topic for future research), outcomes might exceed those that result if indirect training is provided (which more often occurs in school settings; Sterling-Turner et al., 2001). The implications from results of the current study not only suggest a relationship between direct intervention training and an increase in IRA between an observer’s and teachers’ treatment adherence ratings, but also provide preliminary data that support the technical adequacy of teacher self-report of treatment integrity. Although replications of the current study and further empirical evaluation of the effects of direct training are needed, results offer an initial step in increasing the efficiency of school-based consultation without compromising the reliability of data collected related to teacher behavior.
There are design and methodological limitations to consider when interpreting the results of the current study. Limitations to the design include threats to internal, external and construct validity. Specifically, selection of participants was not random. Teachers volunteered, potentially threatening internal validity with possibility of a selection bias. In other words, teachers who volunteered might be more receptive to training than other teachers. Furthermore, with only three teacher participants, generalizability of results might be limited. The teachers taught upper elementary school math and middle school science in an urban school district. Generalizing results to teachers in other educational settings, content areas, or to the population of all teachers may not be possible.

An additional threat to external validity concerns order effects. All teacher participants received direct intervention training first and responded to the intervention favorably. Therefore, implementation of direct assessment training was not warranted. Subsequent studies might introduce direct assessment training first or conduct a study that runs two concurrent multiple baselines that counterbalances phases for the purpose of comparing participants’ responsiveness to direct assessment training. Additionally, a monomethod bias may affect construct validity, as the Teacher Daily Instruction Form was the only measure used by teachers and observers throughout the study, and teachers were told very little about how to use the form. They did not receive direct assessment training and were not aware that observers only made ratings of 0, 3, 5, 8 and 10 based on if 0, 1, 2 3, or 4 steps within a stage were observed. As such, it is possible that discrepancies between teachers’ and the observer’s ratings were exaggerated. Future research might address this methodological concern and provide a bit more guidance to teachers about how to use such a form.
Furthermore, despite efforts to secure consensus with experts in explicit instruction prior to the primary observer rating teachers’ behavior, as well as the completion of IOA during nearly 30% of observations with a secondary observer, construct validity might still be compromised as the ratings of the primary observer were the only ratings used to measure IRA with teachers (i.e., monorater bias). It should also be noted that the primary observer (i.e., student investigator) was not blind to research questions and hypotheses, compromising experimental control.

Current research suggests that multiple dimensions of treatment integrity may be measured when an intervention is implemented (Sanetti & Kratochwill, 2009a). The current study only addressed adherence to intervention steps, limiting conclusions that might be drawn regarding other dimensions such as the quality of intervention delivery or students’ responsiveness to the intervention, for example. Subsequent research might measure additional or other dimensions of treatment integrity (e.g., quality) and how direct training impacts delivery of intervention components as well as implementers’ ratings of their behavior associated with subsequent dimensions.

Limitations to the method include challenges associated with recruitment, scheduling disruptions, and minimal data collected for analyses. Difficulty recruiting districts, principals and teachers limited the scope of the study to assessing three teachers using one multiple baseline design. With more participants, a second multiple baseline design might have been possible to assess direct assessment training first. Additionally, frequent scheduling disruptions including state testing and school vacations resulted in gaps in data collection. This impacted the total number of observations collected, affecting the reliability of conclusions that can be drawn from visual analysis and effect size calculations. It is also a consideration of the interpretation of kappa coefficients.
presented in the results. Moreover, in adapting the URP-I to assess social validity items were changed. This means that alpha values reported for the original measure are no longer relevant for evaluating the measure used in the current study. Finally, student outcome data was not included in the current study. Without student outcome data, discussion about whether explicit instruction had an effect on students’ academic and behavioral outcomes is not possible. Future research might include the collection of student outcome data to generate conclusions about the effect of explicit instruction on these potential dependent variables.

Conclusions

The goal of this study was to determine if direct intervention and direct assessment training would improve IRA between an observer’s ratings and teachers’ self-report ratings of treatment adherence to explicit instruction. Results from the current study lend support to the possibility that, with direct intervention training, IRA improves between raters. Therefore teacher self-report may produce more reliable data than previously shown in research, and as such, teachers may not need direct assessment training when direct intervention training is offered. As hypothesized, the three teachers overestimated their adherence to the intervention as compared to ratings made by the primary observer. However, for two teachers, the direct intervention training led to improved implementation of explicit instruction. This resulted in the primary observer providing higher ratings that matched the teachers’ ratings more closely (aligning with the hypothesis). Alternatively, the third teacher adjusted her ratings rather than her implementation of explicit instruction, which also resulted in improved IRA with the observer. Despite limitations to the design and method, results from this study may
inform implications that improve treatment integrity assessment, school-based consultation and classroom instructional practice.
References


Briesch, A. M., Chafouleas, S. M., & Riley-Tillman, T. C. (2010). Generalizability and dependability of behavior assessment methods to estimate academic engagement:


TABLES
### Effect Size Calculations

<table>
<thead>
<tr>
<th></th>
<th>PND</th>
<th>SMD  $(M_{DIT} - M_{IT}) \div SD_{IT}$</th>
<th>IRD  $IR_{DIT} - IR_{IT}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Teacher C</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advance organizer</td>
<td>1.00**</td>
<td>N/A</td>
<td>1**</td>
</tr>
<tr>
<td>Student enlistment</td>
<td>1.00**</td>
<td>-1.06**</td>
<td>-.40</td>
</tr>
<tr>
<td>Post-organizer</td>
<td>.70*</td>
<td>-1.84**</td>
<td>-.10</td>
</tr>
<tr>
<td><strong>Teacher B</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advance organizer</td>
<td>1.00**</td>
<td>-.88*</td>
<td>-.75</td>
</tr>
<tr>
<td>Student enlistment</td>
<td>.60*</td>
<td>-1.11**</td>
<td>-.15</td>
</tr>
<tr>
<td>Post-organizer</td>
<td>.80*</td>
<td>-1.29**</td>
<td>-.20</td>
</tr>
<tr>
<td><strong>Teacher A</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advance organizer</td>
<td>.00</td>
<td>-.97**</td>
<td>-.08</td>
</tr>
<tr>
<td>Student enlistment</td>
<td>.00</td>
<td>-1.26**</td>
<td>-.54</td>
</tr>
<tr>
<td>Post-organizer</td>
<td>.00</td>
<td>-1.75**</td>
<td>.53*</td>
</tr>
</tbody>
</table>

**Note.** **indicates large effect size; * indicates moderate effect size. For PND, >.90 is very effective, .70 -.90 is effective, .50 -.70 is questionable and <.50 is ineffective (Scruggs & Mastropieri, 1998). For SMD, an effect of .20 -.50 is small, .50 -.80 is medium and greater than .80 is large (Cohen, 1988). For IRD, an effect of >.50 is small, .50 -.70 is moderate and above .70 is large (Parker, Vannest & Brown, 2009).
Table 2.

*Cohen’s Kappa (k) and Kappa Maximum (k_max) Calculations*

<table>
<thead>
<tr>
<th></th>
<th>Teacher C</th>
<th>Teacher B</th>
<th>Teacher A</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Indirect training (Phase B)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( k )</td>
<td>-.04</td>
<td>-.03</td>
<td>.01</td>
</tr>
<tr>
<td>( k_{\text{max}} )</td>
<td>.10</td>
<td>.10</td>
<td>.37</td>
</tr>
<tr>
<td>( k / k_{\text{max}} )</td>
<td>-.41</td>
<td>-.30</td>
<td>.03</td>
</tr>
<tr>
<td>( n )</td>
<td>15</td>
<td>24</td>
<td>39</td>
</tr>
<tr>
<td><strong>Direct intervention training (Phase C)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( k )</td>
<td>.31</td>
<td>.34</td>
<td>.78</td>
</tr>
<tr>
<td>( k_{\text{max}} )</td>
<td>.36</td>
<td>.46</td>
<td>.78</td>
</tr>
<tr>
<td>( k / k_{\text{max}} )</td>
<td>.86</td>
<td>.74</td>
<td>1.00</td>
</tr>
<tr>
<td>( n )</td>
<td>30</td>
<td>30</td>
<td>18</td>
</tr>
</tbody>
</table>

*Note.* \( n \) is the number of pairs of scores. Each observation included three pairs of scores: one for the advance organizer, one for the student enlistment, and one for the post-organizer stage. Umesh, Peterson, and Sauber (1989) suggested taking the ratio of kappa to kappa maximum to determine the observed proportional agreement between raters.
Table 3.

*Teachers’ Average Adherence to Steps of Explicit Instruction Stages by Intervention Phase*

<table>
<thead>
<tr>
<th>Teacher</th>
<th>Baseline</th>
<th>Indirect Training</th>
<th>Direct Intervention Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advance Organizer</td>
<td>4.20</td>
<td>5.00</td>
<td>10.0</td>
</tr>
<tr>
<td>Student Enlistment</td>
<td>8.40</td>
<td>9.20</td>
<td>9.20</td>
</tr>
<tr>
<td>Post-Organizer</td>
<td>4.20</td>
<td>4.60</td>
<td>7.60</td>
</tr>
<tr>
<td>Teacher B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advance Organizer</td>
<td>5.60</td>
<td>5.25</td>
<td>7.30</td>
</tr>
<tr>
<td>Student Enlistment</td>
<td>6.40</td>
<td>6.88</td>
<td>7.30</td>
</tr>
<tr>
<td>Post-Organizer</td>
<td>3.80</td>
<td>6.10</td>
<td>5.80</td>
</tr>
<tr>
<td>Teacher A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advance Organizer</td>
<td>3.60</td>
<td>4.69</td>
<td>10.0</td>
</tr>
<tr>
<td>Student Enlistment</td>
<td>7.80</td>
<td>9.23</td>
<td>9.67</td>
</tr>
<tr>
<td>Post-Organizer</td>
<td>1.80</td>
<td>5.08</td>
<td>8.00</td>
</tr>
</tbody>
</table>
Table 4.

*Inter-observer Agreement for Primary and Secondary Observers’ Ratings*

<table>
<thead>
<tr>
<th>Date</th>
<th>Teacher A (N = 6)</th>
<th>Teacher B (N = 7)</th>
<th>Teacher C (N = 7)</th>
<th>Observation Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/24/12</td>
<td>100%</td>
<td>75.00%</td>
<td>91.7%</td>
<td>88.9%</td>
</tr>
<tr>
<td>1/31/12</td>
<td>91.7%</td>
<td>91.7%</td>
<td>100%</td>
<td>94.4%</td>
</tr>
<tr>
<td>2/14/12</td>
<td>91.7%</td>
<td>100.00%</td>
<td>91.7%</td>
<td>94.4%</td>
</tr>
<tr>
<td>3/6/12</td>
<td>100%</td>
<td>91.7%</td>
<td></td>
<td>95.4%</td>
</tr>
<tr>
<td>3/14/12</td>
<td></td>
<td></td>
<td>91.7%</td>
<td></td>
</tr>
<tr>
<td>3/16/12</td>
<td>100%</td>
<td>83.3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/19/12</td>
<td></td>
<td></td>
<td>91.7%</td>
<td></td>
</tr>
<tr>
<td>3/22/12</td>
<td></td>
<td>91.7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/27/12</td>
<td>100%</td>
<td></td>
<td>83.3%</td>
<td></td>
</tr>
<tr>
<td>4/3/12</td>
<td></td>
<td></td>
<td>91.7%</td>
<td></td>
</tr>
<tr>
<td>Teacher Average</td>
<td>97.2%</td>
<td>89.3%</td>
<td>91.7%</td>
<td>92.6%</td>
</tr>
</tbody>
</table>

*Note.* Blank cells indicate that an observation by the secondary observer did not occur.
Table 5.

Results from Usage Rating Profile – Intervention¹

<table>
<thead>
<tr>
<th>Explicit Instruction Items</th>
<th>Teacher A</th>
<th>Teacher B</th>
<th>Teacher C</th>
</tr>
</thead>
<tbody>
<tr>
<td>I liked the procedures used in explicit instruction.</td>
<td>6</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>I have the skills needed to implement explicit instruction.</td>
<td>6</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>The amount of time required to implement explicit instruction is reasonable.</td>
<td>6</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>I would need consultative support to implement explicit instruction in the future.</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>I would be excited to use explicit instruction in the future.</td>
<td>6</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>I would have no idea how to implement explicit instruction.</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Explicit instruction could be implemented as frequently as needed (e.g., every day).</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>I could implement explicit instruction by myself.</td>
<td>6</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>I would not be interested in implementing explicit instruction in the future.</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>I understand how to implement explicit instruction.</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>All pieces of explicit instruction could be implemented precisely.</td>
<td>6</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Implementation of explicit instruction would require support from my co-workers.</td>
<td>5</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

Direct Intervention Training Items

<table>
<thead>
<tr>
<th>Direct Intervention Training Items</th>
<th>Teacher A</th>
<th>Teacher B</th>
<th>Teacher C</th>
</tr>
</thead>
<tbody>
<tr>
<td>I liked the procedures used in direct intervention training.</td>
<td>6</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>I have the skills needed to receive direct intervention training.</td>
<td>6</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>The amount of time required to receive direct intervention training was reasonable.</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>I would need consultative support to receive direct intervention training again.</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>I would not be interested in receiving direct intervention training again.</td>
<td>1</td>
<td>5</td>
<td>1</td>
</tr>
</tbody>
</table>

Direct Assessment Training Items

<table>
<thead>
<tr>
<th>Direct Assessment Training Items</th>
<th>Teacher A</th>
<th>Teacher B</th>
<th>Teacher C</th>
</tr>
</thead>
<tbody>
<tr>
<td>I liked the procedures used in direct assessment training.</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>I have the skills needed to receive direct assessment training.</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>The amount of time required to receive direct assessment training was reasonable.</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>I would need consultative support to receive direct assessment training again.</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>I would not be interested in receiving direct assessment training again.</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Note. ¹Chafouleas, Briesch, Riley-Tillman and McCoach (2009)
FIGURES
Figure 1. Ratings on DBR scale for Advance Organizer stage
Figure 2. Ratings on DBR scale for Student Enlistment stage
Figure 3. Ratings on DBR scale for Post-Organizer stage
Figure 4. Difference Between Teachers’ and Observer’s Ratings for Advance Organizer stage
Figure 5. Difference Between Teachers’ and Observer’s Ratings for Student Enlistment stage
Figure 6. Difference Between Teachers’ and Observer’s Ratings for Post-Organizer stage
Appendix A

Decision Tree for Determining Adequate Agreement

Has adequate agreement between the teacher and primary observer been met?

Are any of the teacher’s ratings (across the 3 scales) more than 2 points away from the observer’s ratings?

If YES: Adequate agreement has not been met.

If NO: Continue to the next step.

Are the majority of the teacher’s ratings (2/3 of ratings) within 1 point of the observer’s ratings?

If YES: Adequate agreement has been met.

If NO: Adequate agreement has not been met.
Appendix B

Stages of the Explicit Instruction Intervention

Advance Organizer Stage
- Review previous learning: Review what the class has covered in previous lessons
- Define the content of the lesson
- Personalize the instructional material by discussing the benefits to individual students and the places it might be used
- State lesson procedures or expectations (i.e., expected outcomes)

Student Enlistment Stage
- Prompt involvement from the students to get them actively thinking about the instructional material
- Check for understanding of student comprehension frequently
- Expand students’ responses (i.e., correct, expand, and shape)
- Engineer students’ success by scaffolding or providing specific praise

Post-organizer Stage
- Review instructional material by going over the critical elements of what was taught
- Give direction about what will come next (i.e., what instruction or activities will occur the next day)
- State expectations for future learning or performance (i.e., long-term instructional or performance goals)
- Personalize the instructional material by discussing the benefits to individual students and the places it might be used
Appendix C

Expanded Table of Explicit Instruction Definitions and Sample Dialogue

<table>
<thead>
<tr>
<th>Stage/Steps</th>
<th>Operational Definition for Use in the Present Study (Adapted from Lenz, Ellis, Scanlon, 1996; Schumaker, 1989)</th>
<th>Sample Dialogue</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Advance Organizer Stage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Review previous learning</strong></td>
<td>Definition: The teacher reviews with students what the class has covered in previous lessons by stating what was previously taught or asking students to remember what was previously taught.</td>
<td>“Yesterday we learned how to multiply using counting cups and counting bears” (Miller, 2009; p. 135).</td>
</tr>
<tr>
<td></td>
<td>Examples:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• The teacher asks students what was taught in the last lesson (e.g., topics, subtopics)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• The teacher states what was taught in the last lesson (e.g., topics, subtopics)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• The teacher asks students what was learned in the last lesson (e.g., conclusions, concepts, ideas, results, consequences)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• The teacher states what was learned in the last lesson (e.g., conclusions, concepts, ideas, results, consequences)</td>
<td></td>
</tr>
<tr>
<td><strong>Define the content</strong></td>
<td>Definition: The teacher defines for students the content of the current lesson by stating what will be taught and/or the goals of the current lesson.</td>
<td>“Today I’m going to tell you about one method for resolving conflicts. I’d like you to take notes. Then you’re going to work in small groups to identify examples of how you might apply this method to real-life events” (Miller, 2009; p. 135).</td>
</tr>
<tr>
<td></td>
<td>Examples:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• The teacher states topics that the lesson will cover</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• The teacher states goals of the lesson (e.g., continue to practice previously learned content, or learn new information)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• The teacher asks students to predict what will be covered in the lesson</td>
<td></td>
</tr>
<tr>
<td><strong>Personalize the instructional material</strong></td>
<td>Definition: The teacher personalizes the instructional material for students by stating the benefits of learning the new material. The teacher might also state or ask students when and where else the instructional material might be used.</td>
<td>“If you learn to generalize [this] strategy to use in your other subjects, then it is more likely that you are going to write better papers and get better grades on those papers... If you had used [this] strategy on the history assignment last night, then what might have happened?” (Lenz et al., 1996; p. 55).</td>
</tr>
<tr>
<td></td>
<td>Examples:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• The teacher states the benefits of learning the instructional material to students</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• The teacher states the times and/or places the instructional material might be used</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• The teacher asks students to predict the benefits of learning the instructional material</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• The teacher asks students to predict other times and/or places the instructional material might be used</td>
<td></td>
</tr>
<tr>
<td>State lesson procedures or expectations</td>
<td>Definition: The teacher states lesson procedures or expected outcomes to students by informing students of the activities in which they will be engaged or what type of performance that is expected.</td>
<td>“We are going to spend about 20 minutes practicing as a group. I will lead the practice and ask you to volunteer to help me out. Everyone will need to get at least three or four chances to help me out. Then you will work independently for the rest of the period. I will be walking around the room to check on you while you are working. Okay, let’s begin” (Lenz et al., 1996; p. 56).</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td><strong>Examples:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• The teacher states activities in which the students will be engaged during the class period (e.g., group work, independent work)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• The teacher states expectations for student performance (e.g., completion of assignment before the bell rings, independent use, generalization, mastery)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Student Enlistment Stage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Prompt involvement</strong></td>
<td>Definition: The teacher prompts involvement from the students in actively thinking about the instructional material.</td>
<td>“Ok, what is the first thing we want to do when writing a business letter?” (Miller, 2009; p. 138).</td>
</tr>
<tr>
<td><strong>Example:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• The teacher asks the students to participate in teaching the instructional material (e.g., asks questions about the lesson)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• The teacher asks specific questions related to the lesson content, use or related information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• The teacher prompts students to respond (pointing at a student who has a question)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Check for understanding</strong></td>
<td>Definition: The teacher checks for students’ understanding of the instructional material being taught.</td>
<td>“What was the first thing we need to do… the second… the third?” (Lenz et al., 1996; p. 61).</td>
</tr>
<tr>
<td><strong>Example:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• The teacher pauses and asks students if there are any questions about the instructional material (passive)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• The teacher asks direct questions about specific components of the strategy (active)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Expand students’ responses</strong></td>
<td>Definition: The teacher expands students’ responses by correcting and/or shaping answers to questions posed to students.</td>
<td>“What have we covered so far… [after students’ response]”... can you provide more detail about that? (Lenz et al., 1996; p. 57).</td>
</tr>
<tr>
<td><strong>Example:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• The teacher asks students questions about the instructional material and corrects or shapes students’ responses</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• The teacher prompts, cues, hints, restates, rephrases questions, gets the students to the right answer and then restates the original question.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Engineer students’ success (e.g.,)</strong></td>
<td>Definition: The teacher engineers students’ success by reinforcing and scaffolding student progress on the task.</td>
<td>“This is very important… You are doing a great job, but let’s go over it again to be sure that...”</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reinforce, scaffold</td>
<td>Examples:</td>
<td>Post-organizer Stage</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----------</td>
<td>---------------------</td>
</tr>
</tbody>
</table>
| • The teacher reinforces students’ success on mastering the instructional material by providing specific praise  
• The teacher scaffolds students’ learning (e.g., breaks instruction down into discrete steps that increase in complexity)  
• The teacher completes the activity with the students accurately. |

**Examples:**

- The teacher reinforces students’ success on mastering the instructional material by providing specific praise.
- The teacher scaffolds students’ learning (e.g., breaks instruction down into discrete steps that increase in complexity).
- The teacher completes the activity with the students accurately.

*we do this correctly]. Be sure to take notes on this. Write this down. Do you think this is something that is important? What are the most important bits of the information for you to remember at this point?“* (Lenz et al., 1996; p. 57).

**Post-organizer Stage**

<table>
<thead>
<tr>
<th>Review instructional material</th>
<th>Definition: The teacher reviews the critical elements of the instructional material.</th>
</tr>
</thead>
</table>
| Examples:                    | • The teacher states the critical elements of the instructional material  
• The teacher asks students to summarize the critical elements of the instructional material verbally or in writing |

*“Let’s wrap up today by reviewing what we learned…”* (Lenz et al., 1996; p. 59).

<table>
<thead>
<tr>
<th>Give direction about what will come next</th>
<th>Definition: The teacher gives directions about what instruction or activities will come next.</th>
</tr>
</thead>
</table>
| Examples:                                | • The teacher states what material will be taught next  
• The teacher states or describes what activities will be completed next  
• The teacher asks students to predict what instruction or activities will come next |

*“Once we take [our quiz] tomorrow, we will start talking about [content for the next lesson]”* (Lenz et al., 1996; p. 60).

<table>
<thead>
<tr>
<th>State expectations for future learning or performance</th>
<th>Definition: The teacher states expectations that students are capable of learning or how students will benefit from learning the instructional material.</th>
</tr>
</thead>
</table>
| Examples:                                              | • The teacher states future expectations for student learning that are appropriately high to achieve mastery  
• The teacher states the benefit of learning the instructional material  
• The teacher asks students to predict why it is beneficial to learn or master instructional material |

*“The [instructional material] is something you have to learn...I know that when you learn [this material well] and combine it with your effort, you are going to be more successful in your classes...”* (Lenz et al., 1996; p. 60).

<table>
<thead>
<tr>
<th>Personalize the instructional material</th>
<th>Definition: The teacher personalizes the instructional material for students by stating the benefits of learning the new material. The teacher might also state or ask students when and where else the instructional material might be used.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Examples:</td>
<td>• The teacher states the benefits of learning the instructional material to students</td>
</tr>
</tbody>
</table>

*“If you learn to generalize [this] strategy to use in your other subjects, then it is more likely that you are going to write better papers and get better grades on those papers... If you had used [this] strategy on the history assignment last night, then what might have
| The teacher states the times and/or places the instructional material might be used |
| The teacher asks students to predict the benefits of learning the instructional material |
| The teacher asks students to predict other times and/or places the instructional material might be used |

happened?” (Lenz et al., 1996; p. 55).
Appendix D

Teacher Daily Instruction Form

Teacher Name: _________________________                                        Date:  _________________
Topic of instruction: _____________________   Approximate time of lesson: ________________
Time now: ____________________

Directions: Immediately after class, please provide a rating on the scales for items 1, 2 and 3. Treat all sub-questions within each item equally.

1. At the start of the lesson, did you:
   - Review previous learning?
   - Define the content?
   - Personalize the instructional material?
   - State lesson procedures or expectations?

   % complete
   Not at All | Somewhat | All

2. When students practices, did you:
   - Prompt involvement?
   - Check for understanding?
   - Expand students’ responses?
   - Engineer students’ success (e.g., reinforce, scaffold)?

   % complete
   Not at All | Somewhat | All

3. At the end of the lesson, did you:
   - Review instructional material?
   - Give direction about what will come next?
   - State expectations for future learning or performance?
   - Personalize the instructional material?

   % complete
   Not at All | Somewhat | All
Appendix E

Social Validity Measure
(Adapted from the Usage Rating Profile-Intervention (URP-I); Chafouleas, Briesch, Riley-Tillman, & McCoach, 2009)

Name: ___________________
Date: ________________
Class Period: __________

I. Please indicate how much you agree with the following questions about using explicit instruction in your classroom:

<table>
<thead>
<tr>
<th></th>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Somewhat Agree</th>
<th>Somewhat Disagree</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I liked the procedures used in explicit instruction.</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>I have the skills needed to implement explicit instruction.</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>The amount of time required to implement explicit instruction is reasonable.</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>I would need consultative support to implement explicit instruction in the future.</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>I would be excited to use explicit instruction in the future.</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>I would have no idea how to implement explicit instruction.</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>Explicit instruction could be implemented as frequently as needed (e.g., every day).</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>I could implement explicit instruction by myself.</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>I would not be interested in implementing explicit instruction in the future.</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>I understand how to implement explicit instruction.</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td>All pieces of explicit instruction could be implemented precisely.</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td>Implementation of explicit instruction would require support from my co-workers.</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>
II. Please indicate how much you agree with the following questions about receiving **direct intervention training** from the student investigator:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Somewhat Agree</th>
<th>Somewhat Disagree</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>I liked the procedures used in direct intervention training.</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>I have the skills needed to receive direct intervention training.</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>The amount of time required to receive direct intervention training was reasonable.</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>I would need consultative support to receive direct intervention training again.</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>17</td>
<td>I would not be interested in receiving direct intervention training again.</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

III. Please indicate how much you agree with the following questions about receiving **direct assessment training** from the student investigator:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Somewhat Agree</th>
<th>Somewhat Disagree</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>I liked the procedures used in direct assessment training.</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>19</td>
<td>I have the skills needed to receive direct assessment training.</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>The amount of time required to receive direct assessment training was reasonable.</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>21</td>
<td>I would need consultative support to receive direct assessment training again.</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>22</td>
<td>I would not be interested in receiving direct assessment training again.</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Please return to the student investigator, Lindsay Fallon. Thank you!
Appendix F

Teacher Information Sheet and Consent Form

University of Connecticut

Principal Investigator: Lisa Sanetti, PhD
Student Researcher: Lindsay Fallon, MA
Study Title: Direct Training to Increase Inter-rater Agreement between an Observer’s and Teachers’ Self-Report Ratings of Treatment Integrity

Introduction

You are invited to participate in a research study about increasing agreement between ratings provided by an observer and a teacher on how completely a teacher uses explicit instruction in his or her classroom. The study will determine if direct training on steps of explicit instruction and/or direct training on a rating tool will increase agreement between ratings made by a teacher and an observer when agreement is low.

Why is this study being done?

The purpose of this research study is to help build support for using teacher self-report as a method for assessing a teacher’s treatment integrity. Treatment integrity is the degree to which an intervention is implemented as planned by a teacher. Currently, educational researchers often observe teachers directly to determine their level of treatment integrity, but this is not always practical in schools to do. Therefore, this study aims to increase the reliability of teachers’ self-report ratings of treatment integrity so that self-report may be a practical, feasible alternative for treatment integrity assessment in schools.

What are the study procedures? What will I be asked to do?

If you agree to take part in this study, you will be asked to do the following:

If you consent to participate, you will meet with the student investigator for approximately 30 minutes before school, after school, during lunch, or during a preparation period to discuss the study in greater detail. You will be told more about the study, including study procedures, and will have the opportunity to ask questions about the study. You will be told that participation is your choice, and that the risk of participating in the study is minimal.

Once data collection begins, you will be observed delivering instruction during a class period of your choice each day. This class period may be one in which you are experiencing disruptive behavior and want to increase students’ academic engagement, but the choice is yours. It should be the same class period throughout the duration of the study. At the end of the class period you have chosen each day, you will be asked to fill out a form called the Teacher Daily Instruction Form in which you will rate how completely you implemented three stages of explicit instruction. The student researcher will also complete this form and collect this form from you every day.
Sometimes, a second graduate student will also observe your instruction. This is so that
she can provide another set of ratings on the Teacher Daily Instruction Form to be sure
that the student investigator is rating as another observer would. If she is unable to attend
the observation on a particular day, your instruction will be audiotaped so that the
secondary observer can listen to your instruction later to make her ratings. (Note that
audiotapes will be labeled only with the date and a code number so that no identifying
information is linked to the audiotape. Also, audiotapes will be securely locked in a
cabinet when not in use by the researchers.)

Ratings made by you and the student investigator will be compared daily to determine if
there is adequate agreement. If it is determined that ratings are not in adequate agreement
over time, you may be offered more training on how to implement explicit instruction
and how to complete the Teacher Daily Instruction Form.

The total time you may spend participating in this project is approximately 8-10 weeks.

What other options are there?
You always have the option not to participate in the study.

What are the risks or inconveniences of the study?
Although the risks associated with participation in the study are minimal, you may
experience low levels of anxiety during your involvement in the study. Furthermore, at
the end of the study, results will be summarized and graphed. These results may be
shared with school administrators in a short meeting, if requested by school
administrators. Additional inconveniences may include time to meet with researchers and
complete the Teacher Daily Instruction Form each day. If you feel uncomfortable at any
point, you may stop any activity at any time, without penalty.

What are the benefits of the study?
Benefits to participating in this study include potentially (a) improving your instructional
practice by using explicit teaching methods, (b) potentially decreasing displays of student
disruptive behavior, (c) potentially increasing student engagement in your classroom, and
(d) making an important contribution to a growing body of research about how to assess
teachers’ treatment integrity.

Will I receive payment for participation? Are there costs to participate?
There are no costs to participate. As an acknowledgement of your time and effort, you
will be provided with a gift card valued at $50 to Amazon.com at the end of the study.

How will my personal information be protected?
Access to all raw data will be limited to the student investigator, a graduate student who
will help with data collection, the principal investigator, and the student investigator’s
doctoral committee. Code numbers will be assigned for all participants and used at all
times on all study documents. A sheet with the code numbers will be kept apart from the
rest of the data and will only be accessed by the student investigator and the principal
investigator. During the study, any data collected will be stored inside a locked file.
cabinet. All data entered into the study’s database will be kept on a password-protected computer and backed up on a secure hard drive that will be locked up when not in use. The student investigator and principal investigator will be the only people allowed access to the data after the study is over. Data will be stored in a locked file cabinet for 3 years inside a locked office in the Department of Educational Psychology at the University of Connecticut. Electronic data will be stored on the external hard drives that will remain in a locked cabinet as well as on a password protected computer for 3 years as results from the study are being analyzed and published.

You should also know that the UConn Institutional Review Board (IRB) and the Office of Research Compliance may inspect study records as part of its auditing program, but these reviews will only focus on the researchers and not on your responses or involvement. The IRB is a group of people who review research studies to protect the rights and welfare of research participants.

Can I stop being in the study and what are my rights?
You do not have to be in this study if you do not want to. If you agree to be in the study, but later change your mind, you may drop out at any time. There are no penalties or consequences of any kind if you decide that you do not want to participate.

Whom do I contact if I have questions about the study?
Take as long as you like before you make a decision. We will be happy to answer any question you have about this study. If you have further questions about this study or if you have a research-related problem, you may contact the student researcher, Lindsay Fallon (646-942-3436; lindsay.beck@uconn.edu) or the principal investigator directly, Dr. Lisa Sanetti (860-486-2747; lisa.sanetti@uconn.edu). If you have any questions concerning your rights as a research participant, you may contact the University of Connecticut Institutional Review Board (IRB) at 860-486-8802.

Documentation of Consent:
I have read this form and decided that I will participate in the project described above. Its general purposes, the particulars of involvement and possible hazards and inconveniences have been explained to my satisfaction. I understand that I can withdraw at any time. My signature also indicates that I have received a copy of this consent form.

Participant Signature: ____________________  Print Name: ____________________  Date: __________

Signature of Person Obtaining Consent: ____________________  Print Name: ____________________  Date: __________

97
## Abbreviated Table of Explicit Instruction Definitions

<table>
<thead>
<tr>
<th>Stage/Steps</th>
<th>Operational Definition for Use in the Present Study (Adapted from Lenz, Ellis, Scanlon, 1996; Schumaker, 1989)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Advance Organizer</strong></td>
<td></td>
</tr>
<tr>
<td>Review previous learning</td>
<td>Definition: The teacher reviews with students what the class has covered in previous lessons by stating what was previously taught.</td>
</tr>
<tr>
<td>Define the content</td>
<td>Definition: The teacher defines for students the content of the current lesson by stating what will be taught and/or the goals of the current lesson.</td>
</tr>
<tr>
<td>Personalize the instructional material</td>
<td>Definition: The teacher personalizes the instructional material for students by stating the benefits of learning the new material. The teacher might also state or ask students when and where else the instructional material might be used.</td>
</tr>
<tr>
<td>State lesson procedures or expectations</td>
<td>Definition: The teacher states lesson procedures or expected outcomes to students by informing students of the activities in which they will be engaged or what type of performance that is expected.</td>
</tr>
<tr>
<td><strong>Student Enlistment Stage</strong></td>
<td></td>
</tr>
<tr>
<td>Prompt involvement</td>
<td>Definition: The teacher prompts involvement from the students in actively thinking about the instructional material.</td>
</tr>
<tr>
<td>Check for understanding</td>
<td>Definition: The teacher checks for students’ understanding of the instructional material being taught.</td>
</tr>
<tr>
<td>Expand students’ responses</td>
<td>Definition: The teacher expands students’ responses by correcting and/or shaping answers to questions posed to students.</td>
</tr>
<tr>
<td>Engineer students’ success (e.g., reinforce, scaffold)</td>
<td>Definition: The teacher engineers students’ success by reinforcing and scaffolding student progress on the task. Students, offering specific praise.</td>
</tr>
<tr>
<td><strong>Post-organizer</strong></td>
<td></td>
</tr>
<tr>
<td>Review instructional material</td>
<td>Definition: The teacher reviews the critical elements of the instructional material.</td>
</tr>
<tr>
<td>Give direction about what will come next</td>
<td>Definition: The teacher gives directions about what instruction or activities will come next.</td>
</tr>
<tr>
<td>State expectations for future learning or performance</td>
<td>Definition: The teacher states expectations that students are capable of learning or how students will benefit from learning the instructional material.</td>
</tr>
<tr>
<td>Personalize the instructional material</td>
<td>Definition: The teacher personalizes the instructional material for students by stating the benefits of learning the new material. The teacher might also state or ask students when and where else the instructional material might be used.</td>
</tr>
</tbody>
</table>
Appendix H

Expert Raters for Primary Observer Training

1. Patricia G. Sampson Graner, Ph.D., Director of Professional Development,
   University of Kansas, Strategic Instruction Model (SIM) Professional Developer

2. Barbara Ehren, Ed.D., Professor & Chair, Speech-Language Pathology,
   University of Central Florida, SIM Professional Developer

3. Nanette Fritschmann, Ph.D. Assistant Professor, Special Education, The College
   of William & Mary, SIM Professional Developer

4. Gail Cheever, Independent Professional Developer (30+ years of experience)
Appendix I

Draft of Indirect Training Script

Estimated time: 15 minutes

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Dialogue</th>
</tr>
</thead>
</table>
| Start the training session                                       | “In this brief training, we will discuss steps of explicit instruction and how to monitor your delivery of explicit instruction daily. I will tell you about a few procedures that are related specifically to the research study. Throughout the training, you will be able to ask questions about what we’ve discussed. Then, the next time your ____ class meets, you will start to monitor your delivery of explicit instruction and report your implementation of explicit instruction steps on a form that I will provide. You’ll follow this procedure daily until we meet again. Shall we begin?”  
  
  *Record answer:*                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| Introduce explicit instruction                                   | “Explicit teaching methods are important to the delivery of high-quality instruction. Explicit instruction is typically teacher-directed, has an academic focus, requires content to be precisely sequenced, and demands that teachers monitor and provide specific corrective feedback to students when needed. Explicit instruction also involves (a) breaking a task down into smaller steps, (b) breaking instruction into phases, (c) administering probes and student-directed questions, (d) providing visual presentation of instructional material, (e) allowing for independent student practice, and (f) allowing for instruction to be individually-paced. Often, explicit instruction includes providing several opportunities to respond to material learned, and offering students time to practice using new skills or interacting with new material. The goal of explicit instruction is to help students become more independent in executing certain tasks.”  
  
  *Provide an overview for steps of the advance organizer stage:*                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| Provide an overview for steps of the advance organizer stage     | “You are going to begin to report your delivery of explicit instruction in your [X] period class. The way that you do this is by tracking your implementation of steps in three stages of explicit instruction when teaching each day [show teacher the list of stages; Appendix B]. With this sequence of explicit instruction, you start each class with an advance organizer phase. This will prepare students for what they will learn that day before asking them to try the task or new material independently. During this stage, you’ll review what the class has
covered in previous lessons, **define** the context of the lesson, **personalize** the instructional material by discussing the benefits to individual students and the places it might be used, then **state** lesson procedures and expected outcomes. The purpose of the advance organizer is to prepare students for the lesson you will teach them that day. Do you have any questions about the advance organizer stage?”

*Record answer:*

*Question(s):* ________________

**Provide an overview for steps of the student enlistment stage**

During class each day, you will ask students to practice what you have taught them. This is where the students try a new task or practice with new material on their own but with your help. In the student enlistment stage, you’ll **prompt involvement** from the students to get them actively thinking about the instructional material that you presented. You’ll then **check students’ understanding** of new instructional material frequently. Ask students questions about the instructional material and **expand responses** (e.g., correct and shape answers). This will help you to **engineer students’ success** by providing specific praise.

“Do you have any questions about the advance organizer stage?”

*Record answer:*

*Question(s):* ________________

**Provide an overview for steps of the post-organizer stage**

At the end of the period comes the post-organizer stage. This wraps-up the day’s activity and prepares students for the next day’s lesson. In the post-organizer stage, you **review** the critical elements of the instructional material and **give directions** about what instruction or activities will come next. Then, you **state expectations** that students are capable of learning and benefiting from the new instructional material. Finally, you’ll also want to **personalize the instructional material** as you did in the advance organizer stage. This will help students connect what they have learned to the benefits associated with using the material students are taught, and discussing with students where and when they might use the material.”

“Do you have any questions about the advance organizer stage?”

*Record answer:*

*Question(s):* ________________
Introduce study procedures

“As I mentioned before, you will begin to implement explicit instruction the next time your [X] period class meets. After you teach each day, you will fill out this form [show teacher the Teacher Daily Instruction Form] as soon as possible. This form prompts you to rate how completely you implemented the three stages that we just went over. As you will see, there is one scale for all of the steps within a stage. Do your best to rate how completely you implemented the entire stage, treating each step equally when you rate yourself overall.”

Briefly describe Direct Behavior Rating

“These 3 scales use Direct Behavior Rating or DBR. DBR is a tool that involves a brief rating of target behavior immediately following a specified observation period. The defining features of DBR are that they are direct (meaning ratings are recorded immediately at the end of an observation) behavior (meaning specific behaviors are rated such as academic engagement and disruptive behavior) ratings (meaning that ratings are conducted repeatedly in a psychometrically sound manner similar to behavior rating scales).

Review Teacher Daily Instruction Form

“As you can see, the 3 questions on this form [point to Teacher Daily Instruction Form] ask you to provide your answer on a DBR scale from 0-10. Again, the scales refer to how completely you implemented each stage of explicit instruction that day. Treating all steps within stage equally, provide a rating from 0-10 on all scales.”

“Do you have any questions about the Teacher Daily Instruction Form?”

Record answer:________

Question(s):____________________________________________

Review study procedures

“Great. So the next time your [X] period class meets, you will track your implementation of explicit instruction. At the end of the class, you will rate the completeness with which you implemented the three stages of explicit instruction using the 3 DBR scales on the Teacher Daily Instruction Form. Do you agree?”

Record answer:________

“Do you have any questions about the procedures we just discussed?”

Record answer:________

Question(s):____________________________________________
Appendix J

**Draft of Direct Intervention Training Script**

Estimated Time: 30-45 minutes

| Start the training session | “Hi, ________. Thank you for meeting with me again! You have done a great job with implementing explicit instruction and filling out the Teacher Daily Instruction Form each day. I wanted to meet with you today to provide you with a little extra training about implementing explicit instruction specifically. I thought we might start with a video that goes over the stages of explicit instruction in more detail than I gave you before. Then we’ll do a few activities. How does that sound?”
| Record answer: __________ |

| Show direct intervention training video | Play direct intervention training video. (Video will model stages of explicit instruction for teachers.) |

| Role-play the advance organizer stage with teacher | “Now that you have viewed the video, I thought we might role-play delivering explicit instruction. Let’s start with the advance organizer stage. Let’s pretend you are teaching a lesson and I am your student. First, what should the topic of our lesson be?”
| Record answer: __________________ |

| Provide positive and corrective feedback for the advance organizer stage with teacher | “Great! Let’s go with that. So, you’ll first introduce the lesson to me using the steps listed here under the advance organizer stage [point to list of stages; Appendix B]. Why don’t you give it a try? I’ll give you feedback after you practice.” |

| Role-play the student enlistment stage with teacher | “That was great! You did many things very well, including [list steps done well]. I think you might be able to work on [list steps in need of improvement and provide suggestions for how to improve].” |

Next, you’ll ask students to practice with the material you’ve taught using the steps listed here under the student enlistment stage [point to list of stages; Appendix B]. Why don’t you give it a try? I’ll give you feedback after you practice.”
<table>
<thead>
<tr>
<th>Provide positive and corrective feedback for the student enlistment stage with teacher</th>
<th>“That was great! You did many things very well, including \textit{list steps done well}. I think you might be able to work on \textit{list steps in need of improvement and provide suggestions for how to improve}.”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Role-play the post-organizer stage with teacher</td>
<td>“Finally, you’ll end the class period by reviewing what you’ve taught that day, making student connections to material and planning for evaluation of student learning using the steps listed here under the post-organizer stage \textit{point to list of stages; Appendix B}. Why don’t you give it a try? I’ll give you feedback after you practice.”</td>
</tr>
<tr>
<td>Provide positive and corrective feedback for the post-organizer stage with teacher</td>
<td>“That was great! You did many things very well, including \textit{list steps done well}. I think you might be able to work on \textit{list steps in need of improvement and provide suggestions for how to improve}.”</td>
</tr>
</tbody>
</table>
| End training | “To review, the next time your \textit{X} period class meets, you will track your implementation of explicit instruction. At the end of the class, you will rate the completeness with which you implemented the three stages of explicit instruction using the 3 DBR scales on the Teacher Daily Instruction Form. I will do the same. We will continue with this procedure until I ask you to meet with me again, which should be in about one to two weeks. Do you have any questions?

\textit{Record answer:} __________

\textit{Question(s):} _______________________________________

___

___

“Thank you for your time today!” |
Appendix K

Direct Intervention Training Slides for Video Presentation

Implementing Explicit Instruction
Lindsay Fallon, MA
University of Connecticut
In this video, you will:

- Be re-introduced to explicit instruction
- Understand the four stages of explicit instruction
- Learn more about each step involved in the four stages of explicit instruction
- Watch demonstrations of teacher modeling the four stages of explicit instruction

What is explicit instruction?

- Explicit teaching methods are important to the delivery of high-quality instruction.

- Explicit instruction is typically teacher-directed, has an academic focus, requires content to be precisely sequenced, and demands that teachers monitor and provide specific corrective feedback to students when needed.
What is explicit instruction?

- Explicit instruction also involves:
  a) breaking a task down into smaller steps
  b) breaking instruction into phases
  c) administering probes and student-directed questions
  d) providing visual presentation of instructional material
  e) allowing for independent student practice
  f) allowing for instruction to be individually-paced.

What is explicit instruction?

- Often, explicit instruction includes **modeling** for students when new material is introduced, providing several **opportunities to respond** to material learned, and offering students time to **practice** using new skills or interacting with new material.

- The goal of explicit instruction is to help students become more independent in executing certain tasks.
What is involved in explicit instruction?

Three parts to explicit instruction:

I. **Advance Organizer** - prepares student to receive information

II. **Student Enlistment** - guided practice with students to ensure understanding

III. **Post Organizer** - evaluate understanding and monitors future progress

Part I: Advance Organizer

Steps involved in the Advance Organizer stage:

* **Review** previous learning

* **Define** the content

* **Personalize** the instructional material

* **State** lesson procedures or expectations
Part II: Student Enlistment

Steps involved in the Student Enlistment stage:

- *Prompts involvement*
- *Checks for understanding*
- *Expands students’ responses*
- *Engineers students’ success* (e.g., reinforce, scaffold)
Part III: Post Organizer

Steps involved in the Post Organizer stage:

*Review* instructional material

*Gives direction* about what will come next

*States expectations* for future learning or performance

*Personalize* the instructional material
**Appendix L**

**Direct Assessment Training Script**

Estimated Time: 30-45 minutes

<table>
<thead>
<tr>
<th>Task</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start the training session</td>
<td>“Hi, [insert teacher’s name]. Thank you for meeting with me again! You have done a great job with implementing explicit instruction and filling out the Teacher Daily Instruction Form each day. I wanted to meet with you today to provide you with a little extra training about how to complete the Teacher Daily Instruction Form specifically. I thought we might start with a video that goes over Direct Behavior Rating, or DBR, and how to make ratings about your behavior using the Teacher Daily Instruction Form in more detail than I gave you before. Then I’ll ask you to do a few activities with me. How does that sound?”</td>
</tr>
<tr>
<td>Show direct assessment training video</td>
<td>Play Part I of direct assessment training video. (Part I includes a review of DBR.)</td>
</tr>
</tbody>
</table>
| Provide instructions on rating practice with examples and non-examples of target behaviors | “Now that you have viewed the video, I’d like us to practice making ratings using the Teacher Daily Instruction Form. I will show you a series of 1-min video clips of explicit instruction. At the end of each clip, you rate the stage of explicit instruction that corresponds with the video clip I showed you on the Teacher Daily Instruction Form. Then I will tell you what rating was derived by asking experts in the field to rate these clips. If your rating does not match the rating given by experts, I will provide you with feedback and you can ask me any questions you have about rating behavior. Then, we’ll watch the clip a second time so that you can see the video clip associated with the rating experts provided. Do you have any questions?”

Record answer: __________

Question(s): ____________________________________________

| Practice with examples and non-examples of target behaviors | Play Part II of direct assessment training video. (Part II includes a series of 1-minute video clips. Each video clip will be followed by an explanation of the “expert rating” and a reviewing of the video, if necessary.) There will be 6 video clips total.

After each clip, say:

“Go ahead and provide a rating based on that clip. [Allow for teacher to make rating]. What was your rating? [Listen for teacher’s rating.] You rated [insert behavior] a [insert rating]. Experts in the field gave the teacher in this clip a [insert rating]. These ratings...

- match! Great job making a rating that matches mine.”
- do not match. [Student investigator provides immediate, corrective feedback]. Let’s watch the video again, keeping in mind the rating experts provided.”

[AFTER the end of Part II of the video, say] “Do you have any questions?”

Record answer: __________

Question(s): ____________________________________________

| End training | “To review, the next time your [X] period class meets, you will track your implementation of explicit instruction. At the end of the class, you will rate the completeness with which you implemented the three stages of explicit instruction using the 3
DBR scales on the Teacher Daily Instruction Form. I will do the same. We will continue with this procedure until I ask you to meet with me again, which should be in about one to two weeks. Do you have any questions?

*Record answer:*

*Question(s):* __________________________

“Thank you for your time today!”
Using the Teacher Daily Instruction Form
Lindsay Fallon, MA
University of Connecticut

Part I
In Part I, you will:

- Be re-introduced to Direct Behavior Rating (DBR)
- Review how to make ratings using the Teacher Daily Instruction Form

Later, in Part II, you will:

- Practice making ratings using the Teacher Daily Instruction Form

Review of Direct Behavior Rating (DBR)

- A tool that involves a brief rating of target behavior immediately following a specified observation period

- Educational Professionals have long used DBR-like tools to provide information to communicate about child behavior (e.g., home school notes, daily behavior report card)
Review of Direct Behavior Rating (DBR)

- The defining features of DBR are that they are
  - Direct
    - Ratings are recorded immediately at the end of an observation
  - Behavior
    - Specific behaviors are rated such as Academic Engagement and Disruptive Behavior
  - Ratings
    - Ratings are conducted repeatedly in a psychometrically sound manner similar to behavior rating scales

Example:
DBR Single-Item Scales (DBR-SIS)

- DBR-SIS is a scale format that has only one target rated per scale. Typically, a single broad behavior (e.g., disruptive) is used to represent a class of behaviors in general (e.g., out of seat, playing with objects).

Academically Engaged

Interpretation: The student displayed academically engaged behavior during 80% of large group math instruction today.
Review how to use Teacher Daily Instruction Form

- When making a rating, think about the entire stage you are rating
  - Did you implement all steps of the stage?
  - If yes, give yourself a 10
  - If not, determine which steps you did not implement. Treat all steps equally and subtract the value for the number of steps missed from 10.

- For example, if you implemented 3 of 4 steps of the advance organizer stage, subtract 2.5 from your rating of 10 to get 7.5. Round up and give yourself an 8.

Video 1: Advance Organizer

Clip 1: Advance Organizer
Advance Organizer

Steps involved in the Advance Organizer stage:

- Review previous learning
- Define the content
- Personalize the instructional material
- State lesson procedures or expectations

1. At the start of the lesson, did you:
   - Review previous learning?
   - Define the content?
   - Personalize the instructional material?
   - State lesson procedures or expectations?

% complete

Not at All | Somewhat | 7 | 8 | 9 | All
-----------|----------|---|---|---|---
Student Enlistment

Steps involved in the Student Enlistment stage:

Prompts involvement
Checks for understanding
Expands students’ responses
Engineers students’ success (e.g., reinforce, scaffold)

2. When students practices, did you:
Prompt involvement?
Check for understanding?
Expand students’ responses?
Engineer students’ success (e.g., reinforce, scaffold)?

% complete

Not at All Somewhat

0 1 2 3 4 5 6 7 8 9 10 All

Video 3: Post Organizer

Raglin
M. Bailey
Anshuldin
Copsey

Kane
Author
Tate
Stone
Post Organizer
Steps involved in the Post Organizer stage:

- Review instructional material
- Gives direction about what will come next
- States expectations for future learning or performance
- Personalize the instructional material

3. At the end of the lesson, did you:
   - Review instructional material?
   - Give direction about what will come next?
   - State expectations for future learning or performance?
   - Personalize the instructional material?

Part II:
Now it’s your turn to practice!
Video 4: Advance Organizer

Advance Organizer

Steps involved in the Advance Organizer stage:

*Review* previous learning
*Define* the content
*Personalize* the instructional material
*State* lesson procedures or expectations

1. At the start of the lesson, did you:
   - Review previous learning?
   - Define the content?
   - Personalize the instructional material?
   - State lesson procedures or expectations?

<table>
<thead>
<tr>
<th>% complete</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Not at All</td>
<td>Somewhat</td>
<td>All</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

121
Video 5: Student Enlistment

Student Enlistment

Steps involved in the Student Enlistment stage:

1. Prompts involvement
2. Checks for understanding
3. Expands students’ responses
4. Engineers students’ success (e.g., reinforce, scaffold)

2. When students practices, did you:
   Prompt involvement?
   Check for understanding?
   Expand students’ responses?
   Engineer students’ success (e.g., reinforce, scaffold)?

% complete

0 1 2 3 4 5 6 7 8 9 10
Not at All Somewhat

122
Video 6: Post Organizer

Post Organizer
Steps involved in the Post Organizer stage:

*Review* instructional material
*Gives direction* about what will come next
*States expectations* for future learning or performance
*Personalize* the instructional material

3. At the end of the lesson, did you:
Review instructional material?
Give direction about what will come next?
State expectations for future learning or performance?
Personalize the instructional material?

% complete

123
Appendix N

Solutions for Deviations from Study Procedures

<table>
<thead>
<tr>
<th>Deviation (General)</th>
<th>Deviation (Specific)</th>
<th>Planned Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absence</td>
<td>Teacher is absent</td>
<td>• No data will be collected for that day</td>
</tr>
<tr>
<td>Non-instructional Class Activities</td>
<td>Exam or quiz given for more than half of the class period (including state testing)</td>
<td>• No data will be collected for that day</td>
</tr>
<tr>
<td></td>
<td>Exam or quiz given for less than half of the class period</td>
<td>• Data collection will proceed as usual</td>
</tr>
<tr>
<td></td>
<td>Movie shown for more than half of the class period</td>
<td>• No data will be collected for that day</td>
</tr>
<tr>
<td></td>
<td>Movie shown for less than half of the class period</td>
<td>• Data collection will proceed as usual</td>
</tr>
<tr>
<td>Schedule Changes</td>
<td>Snow day</td>
<td>• No data will be collected for that day</td>
</tr>
<tr>
<td></td>
<td>Shortened period due to snow delay or half day schedule</td>
<td>• Data collection will proceed as usual</td>
</tr>
<tr>
<td></td>
<td>Assembly for the duration of the class period</td>
<td>• No data will be collected for that day</td>
</tr>
<tr>
<td></td>
<td>Fire drill that results in less than half of the class period missed</td>
<td>• Data collection will proceed as usual</td>
</tr>
<tr>
<td></td>
<td>Fire drill that results in more than half of the class period missed</td>
<td>• No data will be collected for that day</td>
</tr>
<tr>
<td>Missing Ratings</td>
<td>Teachers do not provide ratings after class</td>
<td>• If the student investigator has not received ratings at the end of the school day, the investigator will prompt the teacher to provide ratings prior to the start of the next class</td>
</tr>
</tbody>
</table>