

EXAMINING THE EFFECTS OF ACADEMIC TEAM-INITIATED PROBLEM
SOLVING PROFESSIONAL DEVELOPMENT ON DATA-BASED
DECISION MAKING FOR READING SUPPORTS

by

PAUL M. MENG

A DISSERTATION

Presented to the Department of Special Education and Clinical Sciences
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

September 2019

DISSERTATION APPROVAL PAGE

Student: Paul M. Meng

Title: Examining the Effects of Academic Team-Initiated Problem Solving Professional Development on Data-based Decision Making for Reading Supports

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Special Education and Clinical Sciences by:

Robert Horner	Chairperson
Kent McIntosh	Core Member
Roland Good, III	Core Member
Gerald Tindal	Institutional Representative

and

Janet Woodruff-Borden	Vice Provost and Dean of the Graduate School
-----------------------	----------------------------------------------

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded September 2019

© 2019 Paul M. Meng

DISSERTATION ABSTRACT

Paul M. Meng

Doctor of Philosophy

Department of Special Education and Clinical Sciences

September 2019

Title: Examining the Effects of Academic Team-Initiated Problem Solving Professional Development on Data-based Decision Making for Reading Supports

A significant knowledge base has been developed within the educational literature on how to effectively use students' reading data to identify students who are at-risk for reading failure and which interventions may be effective in supporting them. Despite this, two-thirds of American fourth graders read below proficiency as reported in findings of the most recent National Assessment of Educational Progress. The literature makes two things quite clear: (a) effective decision rubrics exist for how to identify which students need extra support and what support they need, and (b) teachers and other school staff overwhelmingly have access to the data necessary to utilize these rubrics. The study reported in this dissertation seeks to contribute to what is known about how to effectively implement the decision-making models which are known to be effective in supporting struggling readers. Leveraging the existing literature on structured decision-making found in the positive behavior interventions and supports literature, this study experimentally tests the effects of a newly adapted professional development in Team-initiated Problem

Solving applied to reading support decisions (AcTIPS), on the decision making quality of a school's data team as indicated by percent of points earned on subscales of the Decision, Observation, Recording and Analysis tool, and on students' literacy outcomes as indicated by EasyCBM risk status. Data from a multiple baseline across skills design indicate that the professional development was successful in changing the decision making behavior of the data team across the three fundamental domains of TIPS performance. The team demonstrated clear, immediate, and consistent changes in their performance of Meeting Foundations, Decision Making, and Solution Implementation and Evaluation.

CURRICULUM VITAE

NAME OF AUTHOR: Paul M. Meng

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene
Central Washington University, Ellensburg
Green River Community College, Auburn

DEGREES AWARDED:

Doctor of Philosophy, Special Education, 2019, University of Oregon
Bachelor of Arts, Psychology, 2008, Central Washington University
Associate of Arts, General Studies, 2005, Green River Community College

AREAS OF SPECIAL INTEREST:

Literacy supports for student with or at-risk for developing reading disabilities
with an emphasis on applications of multi-tiered systems of support in
communities containing high rates of low-income households
Tiers II and III positive behavior supports with an emphasis on applications in
communities containing high rates of low-income households
Educational Assessment with an emphasis on efficient measurement of reading
competence for applications to data-based decision making within multi-
tiered systems of support

PROFESSIONAL EXPERIENCE:

Assistant Professor, University of Hawai'i at Manoa, 08/19 – Present

Instructor, Morningside Teachers' Academy, 01/18 – Present

Research Assistant, Educational and Community Supports, 02/18 – 06/19

Graduate Employee, University of Oregon, 09/14 – 06/17

Classroom Teacher, Morningside Academy, 08/10 – 08/12

Behavior Intervention Specialist, The Children’s Village, 09/06 – 07/10

Substitute Teacher, Wahluke School District, 09/08 – 06/09

GRANTS, AWARDS, AND HONORS:

Doctoral Student Scholar, ENLIST, University of Oregon, 2018 – 2019

Graduate Student Scholar, MTA Scholars, Morningside Summer School Institute, 2010

Fellow, Marit Thomas Rhoads Fellowship, Central Washington University, 2009 – 2010

PUBLICATIONS:

Scalise, K., Irvin, P. S., Alresheed, F., Zvoch, K., Yim, H., Park, S., Landis, B., Meng, P. M., Kleinfelder, B., Halladay, L., & Partsafas, A. (2018). Accommodations in digital interactive STEM assessment tasks: Current accommodations and promising practices for enhancing accessibility for students with disabilities. *Journal of Special Education Technology*. doi: 10.1177/0162643418759340

Meng, P. M., McIntosh, K., Classen, J. & Hoselton, R. (2016, February). Does implementation of SWPBIS enhance sustainability of specific programs, such as Playworks? *PBIS evaluation brief*. Eugene, OR: OSEP National Technical Assistance Center on Positive Behavioral Interventions and Supports.

Knowles, C., Meng, P., & Machalicek, W. (2015). Task sequencing for students with emotional and behavioral disorders: A systematic review. *Behavior Modification*, 39(1), 136-166. doi: 10.1177/0145445514559927

Bastable, E., Meng, P. M., & McIntosh, K. M. (under review). Using mixed methods to improve the acceptability and feasibility of an intervention to increase equity in school discipline.

Nese, R. T., Meng, P. M., Breiner, S., & Chaparro, E. (under review). Using stakeholder feedback to improve online professional development opportunities.

ACKNOWLEDGMENTS

Throughout the writing of this dissertation I received a tremendous amount of support and assistance. I would like to express my deepest appreciation to my advisor, Rob Horner, for all of the invaluable feedback and guidance he provided throughout the design of this study, the writing of this manuscript, and my course of study at UO. Also for the meaningful contributions he has made to my thinking on single case designs and data-based decision making more broadly. I would also like to extend my deepest gratitude to the members of my committee. Kent McIntosh, who greatly influenced the way I think about implementation of educational systems and the integration of systems across domains. Roland Good, who has profoundly influenced my thinking on early literacy and measurement. And Gerald Tindal, for his timely and insightful feedback on this project. I would further like to express my appreciation to Anne Todd for sharing her valuable insights on the TIPS model. Finally, I am grateful for the assistance of my colleagues Sean, Becky, and Stephanie for their contributions to the data collection required for this study.

I dedicate this dissertation to my family who have so profoundly influenced me on my journey to this milestone. To my parents, sisters, grandparents, aunts, uncles, and cousins for the love, support, and guidance over the years becoming the person I am today. Finally, to my wife Emily and sons Sage and Cedar. You three were the most critical support and driving motivation to contribute the best that I could. Your love kept me focused and my days bright.

TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION.....	1
Statement of the Problem	1
II. REVIEW OF THE LITERATURE.....	8
Evidence-based Practices in Reading.....	10
Response to Intervention	12
Screening and Progress Monitoring within RTI.....	19
RTI at Scale	23
Decision Making.....	25
Decision Making within Schools.....	28
Team-initiated Problem Solving.....	34
The Present Study.....	39

Chapter	Page
III. METHOD	41
Participants	41
Setting.....	43
Measures	45
Decision, Observation, Recording and Analysis	45
Foundations	46
Decision Making.....	46
Implementation and Evaluation.....	47
Student Behavior	49
Social Validity	49
Academic TIPS (AcTIPS) Training.....	50
Session 1	51
Session 2.....	55
Session 3	60
Coaching.....	61
Session 1	64
Session 2.....	64

Chapter	Page
Session 3	65
Design and Analysis	65
Effect Size.....	67
IV. RESULTS.....	69
Direct Observation Data	69
Student Literacy Outcomes	74
Perceptions of Social Validity	75
V. DISCUSSION.....	77
Summary of Findings	77
Team-Based Decision Making	78
Student Outcomes.....	81
Social Validity	82
Research Design	83
Limitations.....	84
Implications and Future Directions	85
APPENDICES	87

Chapter	Page
A. AcTIPS Professional Development	87
B. AcTIPS Training Manual.....	88
C. Decision, Observation, Recording and Analysis II.....	89
D. Primary Intervention Rating Scale.....	90
E. Adapted Self-Assessment of Contextual Fit	91
REFERENCES CITED	92

LIST OF FIGURES

Figure	Page
1. DORA-II Scores Before and After AcTIPS Training.....	73

LIST OF TABLES

Table	Page
1. Proportion of Students at Elevated Risk of Reading Failure.....	75
2. Perceptions of the Social Validity of AcTIPS Training and TIPS Implementation	76

CHAPTER I

INTRODUCTION

Statement of the Problem

Effective use of early literacy instruction is dependent on individual teachers, and school teams, using student data to problem solve adaptations to instruction and support (Coyne, Kame'enui, & Simmons, 2004; Gersten et al., 2009). While data-based decision-making (DBDM) has been a hallmark of early literacy, curriculum based measurement (CBM), and response to intervention (RtI), only recently has DBDM been extended to broader tasks faced by school teams delivering multi-tiered systems (MTSS) of academic and behavior support (McIntosh & Goodman, 2016). The present research focuses on the Team-Initiated Problem Solving (TIPS) model for team problem-solving. The TIPS approach has been demonstrated to improve the decision-making, solution implementation, and improvement in student outcomes for school teams addressing behavior support challenges (Horner et al., 2018a; Newton, Algozzine, Algozzine, Horner, & Todd, 2011; Newton, Horner, Algozzine, Todd, & Algozzine, 2012a; Todd et al., 2011). The approach has not been formally tested by teams focused on academic problem solving, and the primary aim of the current research will be to determine if a school team focused on early literacy can adopt TIPS procedures with fidelity, and implement TIPS procedures with impact.

Reading achievement is one of the strongest predictors of subsequent academic and career success (Butler, Marsh, Sheppard, & Sheppard, 1985; Stainthorp & Hughes, 2004). Approximately two-thirds of children in the U.S. still perform below proficiency

on end of year summative tests (Bandeira de Mello, Bohrnstedt, Blankenship, & Sherman, 2015). A long history of research has examined both the process of reading (Cattell, 1886) and the most effective methods for teaching children to read (Huey, 1908). With the proliferation of new knowledge and the establishment of the International Reading Association in 1956, reading research emerged as a field unto itself around the middle of the 20th century (Flesch, 1955; Jerrolds, 1977). The field continued to create valuable knowledge over the next 50 years, until the critical mass necessary was for a large-scale synthesis was achieved and executed by the National Reading Panel (NICHD, 2000). Over the past twenty years since the National Reading Panel (2000) released its report synthesizing the existing literature on what aspects and methods of reading instruction are most effective, there has been a tremendous volume of additional research produced leveraging these findings (Balu et al., 2015; Foorman et al., 2016; Gersten et al., 2009; Shanahan et al., 2010). Over this period, significant advances have been made regarding the development of models for improving the schoolwide delivery of effective reading instruction and supports (Baker et al., 2011), responding to the needs of diverse learners (Fuchs & Fuchs, 2007), and developing and validating research-based curricula (Stockard, Wood, Coughlin, & Rasplica Khoury, 2018). Many effective methods for remediating students' various reading difficulties have been validated in the literature (Foorman et al., 2016; Gersten et al., 2009; Shanahan et al., 2010). Additionally, a large number of meta-analyses and syntheses have been conducted examining the relative effectiveness of various educational practices across domains (Hattie, 2008). Hattie's (2008) collection of meta-analyses makes clear that one of the most effective educational practices across domains, is data-based decision making (DBDM). Data-based decision

making allows educators to precisely identify the current performance and needs of each student within a school on the basis of empirically obtained information. Application of DBDM within the context of literacy supports has four core functions: (a) screening, (b) progress monitoring, (c) diagnosis of challenges, and (d) summative programmatic evaluation. However, recent analysis of the effectiveness of this practice within the domain of literacy has produced lower than expected effects (Filderman, Toste, Didion, Peng, & Clemens, 2018). Numerous factors may be at play in the discrepancy between the effects reported in Hattie's synthesis and the results of the more recent work of Filderman and colleagues, but several of the factors highlighted by Filderman et al., warrant special consideration.

When evaluating the effectiveness of data-based decision making applied to students' literacy performance, the first point of note is that use of DBDM is generally effective at improving reading outcomes. Reports of lesser effectiveness may be due to the limited sample size included in Filderman et al.'s (2018) analysis. There were 15 studies that met inclusion criteria, of these 9 compared DBDM with business as usual (BAU), as opposed to a purer test of DBDM involving the same intervention/curriculum with and without DBDM (and thus were not included in the full set of analyses). This left only six studies with which to compare the effects of DBDM applied to reading. Further, while the Filderman et al., synthesis focused on reading it did not focus on any specific reading skill or subset of reading skills. Thus, the total range of skills for which DBDM was applied was quite large relative to the number of studies available for comparison. This results in a sample size which is insufficient to pull out the effects of DBDM by skill interactions. Such interactions must be considered likely given both the relative

literatures pertaining to interventions associated with code-based interventions versus those for meaning-related interventions, and the range of cognitive skills associated with each of these broad sub-domains of reading performance. The clearest points which may be derived from the work of Hattie (2008) and Filderman et al., (2018) are that: (a) DBDM is effective at improving students' reading progress over time, and (b) there is still much needed research in this area concerning how to best implement DBDM with reading data to positively impact student progress. It is further clear that a wide array of DBDM practices were included across both meta-analyses noted, indicating that while response to intervention (RTI) enjoys widespread implementation (Balu et al., 2015), the DBDM processes associated with it, have yet to be codified into precisely defined practices with clearly interpretable instructional implications. The literature on RTI clearly stipulates the application of DBDM but is less clear on the precise process by which this DBDM is best implemented.

One key area of known importance for the effectiveness of DBDM on student outcomes are the systems that support DBDM implementation (Horner et al., 2018a; Newton et al., 2012a; Todd et al., 2011). Of particular relevance to this discussion is the practice of team-based decision making made prevalent by the expansion of multi-tiered systems of support (MTSS). Within the context of a team it is critical that the basic foundations of effective team meetings are in place, each member understands clearly their role and responsibilities, a structured and predictable process guides problem-solving discussions, and that a research-based decision rubric exists to link clearly defined problems with their logically-related solutions (Coyne et al., 2004; Horner et al., 2018a; Newton et al., 2012a; Todd et al., 2011). Further, it is critical that teams engage an

empirically sound problem-solving process comprised of distinct phases related to: (a) problem identification, (b) solution development, (c) solution implementation, and (d) summative evaluation (Deno, 1985; Todd et al., 2011). Hoffman, Jenkins, and Dunlap (2009) conducted a study surveying teachers on their access to reading curriculum-based measurement (R-CBM) data and their use of these data for instructional decision making. The results of this study are clear: teachers have access to the data but do not typically make use of these data for instructional decision making. This is inline with one of the speculations made by Filderman and colleagues regarding the curious prevalence of mastery measures in the set of studies analyzed in their meta-analysis: while R-CBM has a preponderance of data validating its utility for DBDM, screening, and progress monitoring, it is more challenging to link these data to instructional decisions than similar data gathered using mastery measures.

Positive behavioral interventions and supports (PBIS) is one area of educational research that has enjoyed particular success using standard progress monitoring measures for effective DBDM (Horner et al., 2009). Within PBIS, office discipline referrals (ODR) are typically used as a metric for screening and progress monitoring at the universal level, with additional metrics, like percent of point sheet points earned in a given period of time, layered on for students who do not respond adequately to universal supports. The most direct analog between DBDM within PBIS and applications to reading data is likely found when schoolwide teams “drill-down” into their data, viewing the data (ODRs) by location, time of day, grade level, etc. This corresponds well with the type of analysis required of academic teams working with reading data. These teams must analyze data by student group, reading skill, and time of year. There is an analogous drill-down process

that these teams must engage with to meaningfully analyze their data. As a group of students or an individual student demonstrates struggle with a particular skill, the team must look deeper into the performance of that student or group on the related sub-skills that may logically produce such struggle.

The intersection of these findings illuminates one potential solution to the national issue of chronic underachievement in the domain of literacy performance: validation and implementation of a codified DBDM process which makes explicit the link between R-CBM data and the instructional decisions that may be logically derived from them. The current study seeks to address this issue by demonstrating the utility of the Team-initiated Problem Solving (TIPS) model of DBDM to R-CBM data. The TIPS model has been validated within the context of universal behavioral supports across several (Horner et al., 2018a; Newton et al., 2012a), but has not yet been experimentally tested for its effect on decision making related to student academic supports. The TIPS model addresses the critical elements necessary to facilitate the efficient implementation of evidence-based reading supports. This model establishes the basic foundations of effective team meetings, clearly articulates a structured and predictable problem-solving process, and asserts the use of a research-based decision rubric related to the domain of application prior to commencing problem-solving activities (Horner et al., 2018a; Newton et al., 2012a; Todd et al., 2011). Further, this model stipulates a six phase problem-solving process: (a) precision problem identification, (b) goal-setting, (c) solution development, (d) solution implementation, (e) monitoring fidelity, and (f) summative evaluation. These components encompass all four of the phases articulated elsewhere in the literature, but with greater specificity (i.e., goal-setting is distinct in this model from both problem

identification and solution development). A formal curriculum for training school teams to use the TIPS process has been developed, validated and made available online (Todd et al., 2011). This staff training curriculum is being used by district, regional and state trainers throughout the U.S., and as of 2018 over 2700 school teams are actively engaged in using TIPS to manage behavior support within their school (Horner et al., 2018a).

The conceptual coherence between the long history of DBDM in literacy and the active use of data for behavioral problem solving in the TIPS approach is promising. The aims of the current research are to (a) adapt TIPS professional development material to data sources and competencies of early literacy, and then (b) formally examine if team training results in a school team conducting meetings that meet TIPS criteria, using effective literacy-based problem solving, and improving the ability of school teams to implement effective literacy interventions.

The following sections provide first a literature review of early literacy instruction, the role of data-based decision-making, and the link between the new TIPS protocol and literacy content. Next the methodology used in the current study is provided, results are detailed, and a discussion of implications and future directions is offered.

CHAPTER II

REVIEW OF THE LITERATURE

Over 60% of fourth graders in the United States read below a proficient level as determined by end of year standardized tests (National Center for Educational Studies, 2016). This reality exists within a broad societal context wherein proficiency as a reader is critical to success in school and life. Reading proficiency has a meaningful impact on academic success across subjects as children advance through school (National Reading Panel, 2000; Snow, Burns, & Griffin, 1998). Children who fail to achieve acceptable levels of proficiency as readers by the end of third grade face elevated risk for a host of adverse outcomes, including: school failure, problem behavior, out-of-school placement, and incarceration, as well as lower rates of employment as adults (Hernandez, 2011; McIntosh et al., 2006; Sum, Khatiwada, McLaughlin, & Palma, 2009). For school failure specifically, a variable centrally associated with each of the others noted, a child who is not a proficient reader in third grade is four times more likely to experience school failure than a child who is a proficient reader in third grade (Hernandez, 2011). Early reading deficits are also associated with later unemployment and incarceration (Sum, et al., 2009). As troubling as these effects are, they are not uncommon. A significant number of children in the United States are at elevated risk for school failure, developing aberrant behavior patterns, out-of-school placements, un/under-employment in adulthood, and future incarceration. Reading deficit is a common and significant source of risk for each of these adverse outcomes.

Reading supports and their delivery in schools and districts nationwide have shifted in recent years. The US government has long wrestled with how to best deliver

instruction and supports to each and every student (Coyne, Kame'enui, & Simmons, 2004; ESSA, 2015; NCLB, 2002). With the widespread adoption of common core state standards, new emphasis has been placed on implementing practices that facilitate students' achievement of a pre-defined criteria. It is worth noting that these standards and the practices supporting their achievement have a long history of both academic research and governmental initiatives that have at times overlapped.

In the recent past, federal expectations under the No Child Left Behind Act of 2002 (NCLB) stipulated that all students must meet state reading proficiency targets by third grade (Coyne, Kame'enui, & Simmons, 2004; NCLB, 2002). Following a complete absence of schools achieving this benchmark by the stipulated 2014-15 deadline, revision to this mandate has been administered and legislated. In 2015, the U.S. Congress passed the Every Student Succeeds Act (ESSA), which had the effect of reversing much of the federal policy established under NCLB. In particular, the new law transferred control of goal-setting from federal to state oversight. Additionally, ESSA made significant changes to the landscape of teacher evaluation and accountability, primarily in removing the requirement that teacher evaluation incorporate student achievement data (ESSA, 2015). Specific federal expectations are now of the form that schools, districts, and states must continue to set goals for improvement and that they do so by relying on evidence-based strategies for improvement.

The National Reading Panel (NRP; 2000) identified five core aspects of early reading instruction: (a) phonological awareness, (b) alphabetic knowledge, (c) vocabulary, (d) fluency with text, and (e) comprehension. These elements constitute the "Big 5 in Early Reading." The NRP report resulted from a comprehensive review of the

literature on reading instruction and development. Recommendations related to specific instructional practices include interventions targeted at phonemic awareness (e.g., blending, segmenting, rhyming), phonics instruction (explicit instruction in letter-sound correspondence), and strategies for building fluency (e.g., guided repeated oral reading). This study has served as the foundational synthesis upon which the majority of reading research conducted since the turn of the century has been built.

A follow-up to the NRP report was conducted approximately ten years later, by a new federally convened panel, examining *early* literacy development and intervention (National Early Literacy Panel [NELP], 2008). The focus of this panel, in contrast to the NRP, was the development of reading and pre-reading competency as it occurs from birth through age five rather than during the early elementary years. Findings from this synthesis indicate that a variety of programs can be used to improve young children's oral language skills. The presumption is that application of these strategies will improve children's early literacy development once exposed to systematic instruction in the early school years, as a function of improved oral language skills.

Evidence-based Practices in Reading

One major policy advancement of the last decade is the special importance that has been placed on interventions that are evidence-based. The What Works Clearinghouse reports five evidence-based elements of support for struggling readers within the framework of RTI including: (a) universal screening, (b) differentiated, data-based instruction, (c) provision of intensive, systematic instruction in small groups on up to 3 targeted reading skills, (d) progress monitoring at least once per month for students at tier 2, and (e) escalation to tier 3 support intensity for those students who make

insufficient progress at tier 2 (Gersten et al., 2008). These elements of RTI implementation focus largely on data-based decision making. The recommendations of Gersten et al. (2008) regarding RTI form the presently most effective framework for delivering evidence-based supports for students struggling learning to read. Within this framework, the evidence-base around reading supports clarifies reasonable intervention and support strategies for learners at different stages of reading acquisition and differing levels of support needs.

For early readers, foundational skills are critical targets that individuals need to learn to develop proficient reading. RTI and the instructional practices associated with this framework reliably result in most students progressing through reading stages at similar chronological ages. Given the importance of proficient reading for future academic success, and the critical importance of getting students on track by the end of third grade, third grade represents a time of particular importance with regard to remediating the skills of struggling readers (Coyne, Kame'enui, & Simmons, 2004). Foorman and colleagues identify four foundational reading skills for readers in this age range including: (a) phonemic awareness, (b) decoding, (c) fluency with connected text, and (d) academic language skills (Foorman et al., 2016). These skills appear to constitute the critical foundation for reading with comprehension in the early elementary years. These are consistent with the skills identified in the NRP (2000) report, but emphasize the skills which are most critical early in development of formal reading skill. Phonemic awareness and decoding track closely with phonemic awareness and alphabetic knowledge, fluency with text is identified in both sets, and academic language skills are a slight variation on vocabulary that emphasizes unknown vocabulary and oral language

skills. Comprehension is notably absent from the Foorman (2016) as formal emphasis on comprehension skills typically waits until later elementary years (fourth and fifth grade). The skills identified in Foorman and colleagues' analysis are expected to provide the basis of strong comprehension skill development. In particular, interventions targeting decoding skills and word reading skills have shown significant promise for preventing and remediating deficits in reading proficiency (Simmons, Kame'enui, Stoolmiller, Coyne, & Harn, 2002). A closer look at the core features and body of research supporting this practice is critical to understanding the current landscape of reading instruction, intervention, and student achievement.

As of this writing, 41 instructional or intervention practices have achieved the criteria set by the What Works Clearinghouse as evidence-based practices for elementary-aged students ([wwc.org](http://www.wwc.org)). Practices in this category range across reading sub-skills but skew heavily in favor of code-based approaches which emphasize decoding skill via phonics and phonemic awareness. These approaches also vary substantially in terms of the ages for which they have been validated. Some including students in preschool through Kindergarten or first grade, others deemed appropriate for a single elementary grade-level (on the basis of established evidence), and others ranging from 5th through 12th grades. Some involve very specific curricula (e.g., Wilson Reading System), while others are broader strategies that can be used with a variety of curricula (e.g., Peer-Assisted Learning Strategies).

Response to Intervention. Response to Intervention (RtI) is a well-studied, evidence-based practice for improving literacy outcomes (Gersten et al., 2008). Three core elements of RtI are its multi-tiered delivery system, data-based decision making, and

utilization of evidence-based instructional practices at each tier. One of the core ideas for RtI within education, adaptive instruction leveraging data-based decision making, traces back to several traditions within the field, including: precision teaching (Lindsley, 1964), curriculum-based measurement (Deno, & Mirkin, 1977), behavioral consultation (Bergan & Kratochwill, 1990; Kratochwill & Bergan, 1978), and Direct Instruction (Engelmann & Carnine, 1982).

Ogden Lindsley brought the first responsive instructional framework based on data to the literature in his early descriptions of precision teaching (Lindsley, 1964). His emphasis was on the application of what the field of behavioral science had learned from the past several decades of work in operant learning laboratories across the country, to the field of special education. Utilizing the methods of measurement germane to the operant laboratory of the 1960s, Lindsley emphasized response frequency as the metric of interest. This focus on rate of responding encouraged a further emphasis on the very small units of learning which combine to create repertoires of academic and social relevance. These very small components, or “pinpoints,” are especially useful in special education where larger skills or concepts often need to be broken down into more easily understood elements (Archer & Hughes, 2011; Engelmann & Carnine, 1982). The emphasis on rate of response also had the advantage of providing numerous practice opportunities as individuals typically complete as many repetitions of the target behavior as they can within a given 1-5 minute period of time (Johnson & Street, 2012). Precision teaching utilizes response rate, as an indicator of the strength of the stimulus control relationship between a target stimulus (i.e., b/a/t) and the appropriate response (i.e., “bat”), to empirically determine when a learner has mastered each piece of a relevant

skill. For each sub-skill, precision teachers establish a goal or “aim” rate which is used as an indicator that a learner has mastered that piece of the larger skill. Once this aim is met, learners advance to the next pinpoint and repeat the process until all sub-skills are mastered and the larger skill can be performed fluently. When learners do not reach the response frequency associated with sub-skill fluency, the precision teacher uses these data to make a determination about what kind of change is needed. An intervention is selected or developed, delivered to the learner, and timed practice resumes- a process that may typically take 3-5 minutes (Johnson & Street, 2012).

Following the demonstration of precision teaching in Lindsley’s KU affiliate research sites, the practice was trialed in more diverse contexts (Binder, 1996; Binder & Watkins, 1990; Datchuk & Kubina, 2017; Johnson & Street, 1996; Lindsley, 1990). These applications demonstrated both the significant promise of the approach for improving student performance and the unique staff costs associated with its implementation. Staff training and time to implement precision teaching as documented in the literature through the early 1990s constitutes a highly intensive intervention (Hayes, Heather, Jones, & Clarke, 2018; Lindsley, 1990). Given the pressing need for efficiency in educational contexts due to limited funding, it thus constitutes an intensive Tier 3 intervention within the context of RTI.

Building on this work, Stan Deno and his colleagues conceptualized their Data-based Program Modification Model of intervention delivery (Deno & Mirkin, 1977). Deno conducted educational assessment work early in his career in Minneapolis Public Schools. This work focused on training teachers in methods for measuring the effects of their instruction on their students’ learning. During this early work, he emphasized the

collection of direct measures of target skills frequently, graphing results, and application of these measures with every student to assess learning within students across time. Deno and Mirkin (1977) articulate the essential logic of this data-based instructional approach. These authors' conceptualization of the data-based decision-making process emphasized the logic of the what decisions were well-supported on the basis of which data. Their model of measurement early on focused on measurement of specific skills and sub-skills that students were working to master, and later transitioned to the broader indicators of proficiency seen in curriculum-based measurement (Deno, 1985). Deno (1985) articulated the potential for curriculum-based measurement as an efficient method for valid and reliable decision making related to screening, referral, programmatic, and progress monitoring decisions, marking a clear shift from the previous practices of informal teacher observation and achievement tests for these purposes. Utilizing materials and procedures that were readily comprehensible for practicing teachers, CBM represented a new level of efficiency in the valid and reliable assessment of student progress and proficiency. Critically, CBM built off of the measurement work within precision teaching by utilizing response rate as the key metric in early measures, thereby leveraging the efficiency of stimulus control measurement for assessing skill proficiency.

Bergan and Kratochwill gave the field a system of data-based decision-making for student social development applied to clinical and school settings with their model of behavioral consultation (Bergan & Kratochwill, 1990; Kratochwill & Bergan, 1978). Their process of decision-making involved four stages: (a) problem identification, (b) problem analysis, (c) intervention, and (d) evaluation. This formulized four-stage approach was a clear precursor to the codified decision-making models which have been

applied subsequently within RTI and multi-tiered systems of support (MTSS) more broadly.

Zig Engelmann, Doug Carnine, and Wes Becker followed with their early work building what is known about explicit instructional methods (Engelmann, 1968; Becker & Engelmann, 1975, 1976, 1978; Engelmann & Carnine, 1982). Engelmann utilized what was known about human learning and skill development from the work being done in operant learning laboratories of the time to develop a generalized method of instruction that was capable of clearly communicating knowledge to diverse groups of students with very little assumed prerequisite knowledge (Engelmann, 1968). In the largest trial comparing instructional approaches in history, Project Follow-through, Direct Instruction (DI) resulted in clearly higher outcomes across measures as compared to all other tested methods, with the exception of the Behavior Analysis model out of the University of Kansas which resulted in performance that was closer to that produced by DI (Becker & Engelmann, 1978; Engelmann, Becker, Carnine, & Gersten, 1988; Stebbins, 1976; Stebbins, Pierre, Proper, Anderson, & Cerva, 1977). Engelmann and Carnine further distilled the core elements of effective, explicit instruction in their seminal text on the topic in 1982 (Engelmann & Carnine, 1982). Since that time, a great deal of research has been conducted validating specific explicit instructional curricula and expanding the theory and procedures associated with this family of instructional methods (Archer & Hughes, 2009; Stockard, Wood, Coughlin, & Khoury, 2018).

Application of the multi-tiered delivery of educational supports was first articulated and brought to significant scale within the positive behavioral interventions and supports literature (Horner et al., 1990; Sugai & Horner, 2002). This model, like RTI,

involves three core levels of preventative support for students based upon indicated need. Primary prevention applies to all students within a school, secondary prevention involves more intensive supports and applies to a smaller subset of students within a school, and tertiary prevention involves the most intensive (individualized) preventative supports and applies to the few remaining students who are not responsive to the previous two levels of support.

Related to its multi-tiered system of support delivery, RtI shares a common lineage with positive behavioral interventions and supports (PBIS; Sugai & Horner, 2002). Tracing the precise origin of the multi-tiered model of service delivery is difficult. Several have noted that public health implemented such a system first (Walker et al, 1996), some tracing back to Caplan and Grunebaum (1967) specifically. However, as others note many variations were in effect and disseminated widely by that time in the medical literature (Gordon, 1983; McIntosh & Goodman, 2016). Within the context of educational applications, Simeonsson (1994) edited the first significant publication describing multi-tiered supports. This text advocated for a broad paradigm shift within education away from a focus on intervention to one of prevention, recognizing the need for escalating supports based upon indicated need. Both RtI and PBIS are part of what is commonly termed multi-tiered systems of support (MTSS) within education. MTSS is a broader classification of systems meeting this core element and several other core features including data-based decision making and reliance on evidence-based practices (see McIntosh & Goodman, 2016).

At present, RtI and MTSS are widely utilized for elementary reading supports within the United States, with some data indicating that over two-thirds of American

public schools are engaged in some level of implementation (GlobalScholar, 2011). Harn et al. (2011) conducted a study looking at the quality of core curricula implemented in practice in two districts within Oregon. Emphasis of core curricula selected by the two districts in terms of targeted subskills, as well as the dosage given to Tier I students, were examined. Further, the degree of alignment across Tiers was carefully analyzed.

Following this initial analysis, the researchers worked with school personnel to intervene, focusing largely on the alignment of supports across Tiers. Alignment was poor or nonexistent at pre-test, limiting the degree to which instruction and practice in supplemental blocks could be applied to content covered during Tier I instruction.

Alignment of curricula across Tiers is observed when the content covered during whole group instruction is related to the content covered in supplemental instructional blocks.

Alignment of instruction across Tiers occurs when the instructional strategies used during supplemental instructional blocks (Tiers 2 and 3) are both appropriate to the learners' level of need (increasingly explicit with smaller units) and complementary to the methods of instruction used during whole group instruction. Following alignment of curricula and instruction across Tiers, students requiring escalated levels of support showed meaningfully improved rates of progress, indicating that alignment of supports was functionally related to positive academic outcomes for these students.

Coyne et al. (2004) conducted a study looking at Kindergarten and first grade intervention. Their results indicate that earlier intervention has a stronger effect and a qualitatively different effect in terms of subsequent rate of growth and risk status. This is consistent with current theory that there is a critical period of early reading development wherein intervention has a stronger preventative effect on subsequent reading difficulty

(Simmons et al., 2008). Further, the research team utilized decision rules for determining assignment to intervention conditions similar to those advocated elsewhere (Deno, 1985; Fuchs & Fuchs 1989; Harn et al., 2011) and found significant effects on student reading outcomes. One critical feature of this study was the application of data-based decision making (DBDM) by expert researchers, not classroom teachers. Taken together, these findings support the idea that the utilization of data is a critical area of concern within the effective implementation of RTI.

Screening and Progress Monitoring within RTI. Screening requires measures which are efficient to administer and sensitive to inter-individual differences in performance at a given point in time (Salvia, Ysseldyke, & Bolt, 2010). Progress monitoring requires measures which are highly efficient to administer (requiring short administration time) and sensitive to intra-individual changes in performance over time (Salvia, Ysseldyke, & Bolt, 2010). Fuchs and Deno (1991) discuss two models of progress monitoring: (a) Mastery Monitoring (MM), and (b) Global Outcome Measurement (GOM). Mastery Monitoring is the practice of frequent measurement of composite skill subskills for student acquisition and mastery. This process involves the frequent (in some models of MM, daily) measurement of student performance on tasks targeting specific subskills. In contrast, the authors define the newly-presented model of GOM for progress monitoring in terms of its focus on long-range goals and its standardization of administration procedures and tested stimuli. The authors make the case that GOM is a far more efficient model of progress monitoring when compared with MM. The authors contend that GOM, the family of progress monitoring measures to which curriculum based measures (CBM) belong, is preferable for progress monitoring

because it: (a) is efficient for teachers to use, (b) has demonstrated validity and reliability, (c) yields information which is useful to instructional planning, and (d) is useful for answering questions of program effectiveness in terms of overall student growth.

Additionally, GOM is sensitive to inter-individual differences in performance at a given point in time, making it useful for both progress monitoring and screening applications (Good & Kaminski, 1996). These four elements were the defining priorities of the CBM model created by Fuchs and Deno. In the process of creating and defining their model of CBM, the authors identify what they consider to be the defining features of GOM more broadly: prescribed procedures for administration and content stimuli, and long-range consistency. The authors make the case that MM, as was prevalently used for progress monitoring in decades past, has two critical failings: (a) excessive flexibility, and (b) a focus on short-term goals. These two failings led to use of ad hoc and idiosyncratic measures by teachers across classrooms, and ineffectual tests for answering questions of student growth over time and comparative evaluations of educational programs and strategies. Indeed, the Mastery Monitoring encountered by the authors lacked systematization. This lack of systematization subsequently resulted in a both a lack of long-range consistency and a lack of applicability to long-term goals. At its core, MM pales in comparison to GOM when the lens of analysis is on the acquisition of long-term goals. This basic limitation of MM is present irrespective of what type of systematization is applied (and especially true given the statistical procedures prevalent and available at the time of this work by Fuchs and Deno).

While GOM is superior in many regards, it is possible that there has been some unintended cost associated with the information that has been lost as a function of the

move entirely away from MM in favor of GOM. In particular, teacher use of data for the purposes of predicting students' response to intervention, and making student progress more salient for teachers, may have been adversely impacted. As Fuchs and Deno note, MM was a commonly used progress monitoring practice *by teachers* at the time of their writing. Teachers' use of progress monitoring data for decision-making is less than common now according to recent data (Hoffman, Jenkins, & Dunlap, 2009). The practices of MM at the time of the previous writing left much to be desired, but the data were collected in terms of activities that teachers used in their teaching, and were thus likely to be readily interpretable. While GOM allows for efficient tracking of student progress over time, it is possible that the data, due to their general nature, are less inherently meaningful for teachers than student performance on classroom learning activities. In order to support teachers' use of these data for effective decision-making, additional training and support systems may be needed. While it is clear that left to their own devices teachers are not using GOM data to inform their instructional decision-making at high rates, to achieve desired levels of student reading proficiency, or sustainably over time, it remains to be seen what strategies or interventions may improve teachers' use or perceptions of these data.

The noted excessive flexibility and emphasis on short-term goals were certainly critical failings of MM as practiced during the 1980s and early 1990s. While relevant to the practice of MM at the time, the lack of prescribed procedures and stimuli within MM is not an essential feature of the model, but rather evidence of insufficient development of measures for the model by the research and development community at that time. While it is certainly true that GOM requires far fewer measures (by at least one order of

magnitude) to operate with prescribed procedures for administration and stimuli than does MM, it is nevertheless entirely possible to utilize MM using only standardized measures. The crux of this issue is the degree to which granular measures appropriate for progress monitoring have been developed within a given broad domain. Within the domain of reading, we may consider the prevalence of oral reading fluency (ORF) within the progress monitoring landscape. While a critical outcome with validity, long-term reliability, and importance, the decisions which can be supported based upon its data will not allow for appropriate instructional modifications for all struggling readers. In fact, while a great many students may avoid abject reading failure through its implementation, the majority of those who struggle will not ever achieve reading success either. This raises a critical question of priority for the field of (special) education more broadly. There is a critical need for more detailed, diagnostic assessment within any system which is intended to serve the needs of struggling students.

As Fuchs and Deno (1991) note, MM is particularly well-suited to answering the question of “did this student learn what I taught today?” This is not among the questions well-answered by CBM or GOM in their present state. However, MM does not answer these questions with reliability or demonstrable validity as practiced at the time of their writing, either. This issue of providing data that is relevant for instructional decision-making is of critical importance when considering the conditions likely to sustain teacher-implementation of any given practice. In the area of implementation science, it is generally accepted that implementer experience with a given practice (how positively a practice is regarded by those charged with implementing it) impacts both fidelity of implementation and sustained implementation over time. Within the domain of teaching

reading specifically, there is some evidence supporting the need to provide data within a time frame and format relevant for classroom decisions (Garet et al., 2008). Practices that do not produce teacher-perceptible changes in student achievement are at risk for abandonment.

RTI at Scale. The essential element of RTI that drives its efficiency is the practice of escalating (diagnostic) assessment and supports for only those students with indicated need. This requires that: (a) student needs are accurately identified, (b) effective interventions are matched to student needs, and (c) interventions are refined over time using progress monitoring data (Coyne, Kame’enui, & Simmons, 2004). However, as a system with multiple critical systems, RtI can be difficult for schools and districts to implement with fidelity (Reynolds & Shaywitz, 2009). Given the centrality of data to the implementation of RTI it is reasonable to hypothesize that this may be an area of critical need. As Hoffman, Jenkins, and Dunlap (2009) note in their study of CBM data use, teachers have access to the data but are not commonly utilizing it for instructional decision making. When data are available but not leveraged to adjust instruction, student needs go unmet.

Further supporting the conclusion that student needs are not being identified with sufficient precision in practice is the corpus of findings suggesting that when highly skilled experts assume the task of matching student needs to interventions, students make much greater gains than under “business-as-usual” or baseline conditions (Harn, Chard, Biancarosa, & Kame’enui, 2011; Simmons, Kame’enui, Stoolmiller, Coyne, & Harn, 2002). The finding that changes in programming made mid-year, determined by expert review of available data, selected from commonly known interventions, further supports

the idea that decision making processes have significant potential for improvement (Simmons, Kame'enui, Stoolmiller, Coyne, & Harn, 2002). Further, data exploring the degree to which teachers utilize data to inform decision making consistently within an RTI framework indicate that this is an area of significant value and insufficient application (Hoffman, Jenkins, & Dunlap, 2009; (Sharp, Sanders, Noltemeyer, Hoffman, & Boone, 2016).

Sharp and colleagues (2016) conducted a study using hierarchical modeling to examine the relationship between RTI implementation and student reading achievement. In their study, they collected data from administrators and school psychologists across 43 elementary schools and controlled for several meaningful demographic characteristics (e.g., school-level socioeconomic status [SES], office disciplinary referrals [ODRs]). Their findings indicate that most schools participating in the study had relatively high levels of Tier I implementation, with relatively small levels of variability across schools on this metric. Their findings also indicated that the demographic variables included accounted for 36% of variance in student reading outcomes. Additionally, subscale scores on the Assessment domain of RTI implementation were the highest of all subscales and demonstrated the lowest degree of variability. The data are generally being collected. Tier III implementation accounted for a significant amount of the variability (6.8% of the variance) in reading outcomes. However, the most critical finding for the present analysis was that implementation of data-based decision making (DBDM) accounted for 7.2% of the variance in reading outcomes, highest among the two modeled predictors and only slightly less than the amount explained by ODRs (8.1%). This finding is critical for at least two reasons. First, exclusionary discipline removes the student from the

instructional environment, and the application of DBDM predicts nearly as much variance in student reading outcomes. Second, in the presence of the finding that assessment is an area of relative strength within RTI implementation for this sample, the relatively low levels of DBDM and the significant amount of variance explained by this variable indicates that it is the utilization of assessment data that is lagging rather than the collection of such data that is most likely impeding full realization of RTI's benefits at scale.

Data-based decision making within RtI comes in several forms based upon different functions of assessment which are built into the model. Two critical functions of decision making which are well integrated into RtI are: (a) screening, and (b) progress monitoring. Different types of assessment are best suited to different functions of assessment; matching a given measure to the function of assessment for which it is best (or at least well-) suited is critical to using the data effectively to make decisions. Screening and progress monitoring are both functions of assessment which require efficient measures in terms of administration time. Screening is best accomplished by measures which identify those who may need additional support as precisely as possible, whereas progress monitoring is best achieved using measures which are highly sensitive to intra-individual changes in performance over time. In some applications, these may be embodied within a single measure. In current practice this is frequently the case, as curriculum-based measures are often used for both screening and progress monitoring. The research on decision making, both broadly and within educational contexts, warrants further inspection.

Decision Making

Research within education and psychology has been conducted on the critical processes involved in making decisions in pursuit of a given goal (D’Zurilla & Goldfried, 1971; Hattie, 2008; Lichtenstein, Fischhoff, & Phillips, 1976). Models have spanned numerous fields and taken differing perspectives on several aspects of decision making, but the critical features of the basic process have been quite consistent across studies over time. Some researchers have emphasized the process of decision making (Nezu, Nezu, & Peri, 1989; Todd et al., 2011), others have emphasized the outcome of decisions (Messick, 1995), while still others have placed emphasis on the data used to guide decisions (Deno, 1985; Fuchs & Fuchs, 2006), while still others have emphasized the use of a particular decision rubric for a specified purpose (Good & Kaminski, 1996). Numerous models of data based decision making have been put forth in the psychological and educational literature (D'zurilla & Goldfried, 1971; Deno, 1985; Fuchs & Fuchs, 2006; Good III & Kaminski, 1996; Nezu, Nezu, & Perri, 1989). Most of these models have identified 4-6 steps for reliable decision making. These steps are generally of the form: (a) identify the problem, (b) propose a solution, (c) test the solution, and (d) evaluate the solution.

Research on decision making in general has revealed several key findings. Chief amongst these is that when individuals make decisions without adequate training they have an overwhelming tendency to demonstrate overconfidence (Lichtenstein et al., 1976). Overconfidence in the context of decision making is when the individual estimating the likelihood of a given outcome consistently offers a probability estimate that exceeds the observed rate of the outcome given the known information at the time of

prediction. In the absence of better information, these subjective probability estimates are demonstrably related to decision making (Lichtenstein et al., 1976). The issue of overconfidence influencing results can be addressed to a significant degree by: (a) additional training and experience with problem-solving with feedback (Donovan, Guss, & Naslund, 2015), operating within a team structure (Gersten et al., 2008), both of these (Todd et al., 2011), or by using an explicit decision rubric in the analysis of pre-specified data (Fuchs & Fuchs, 2006; Good & Kaminski, 1996).

Lichtenstein and colleagues (1976) addressed the issue of overconfidence and offered a solution. In their report, the authors define overconfidence as the observed overestimate of probability for a given event as compared to the observed rate of that event's occurrence across multiple trials. They offer the solution of calibration, the calculation of the exact rate of occurrence across multiple trials, with subsequent comparison to the probability estimated a priori. They offer the technology of calibration as a method for improving the accuracy of decision making in practice, and for refining models of decision making more generally.

D'Zurilla and Goldfried (1971) presented a model of problem solving comprised of five steps: (a) general orientation, (b) problem definition, (c) generation of alternatives, (d) decision making, and (e) verification. They describe problem solving as a behavioral process, functioning to increase the probability of identifying a successful solution. They further stipulate that the process of problem solving facilitates access to a number of alternatives. The authors propose that successful problem solving is a repertoire typically learned through trial-and-error, with successful resolution serving to maintain the performance that produced it.

Nearly two decades later, Nezu and Perri (1989) conducted a study examining the effectiveness of problem-solving as a form of cognitive-behavioral therapy for depression. The authors report that problem-solving was effective in reducing the number of participants experiencing clinically significant depression. Special emphasis was placed on the problem orientation portion of their problem-solving model. Findings indicate both that their model of problem solving was effective in remediating depressive symptoms, and that particular attention to individuals' problem solving orientation is warranted as it is meaningfully associated with subsequent problem-solving success.

Decision Making Within Schools. Formalized problem-solving within schools began to take shape later, but has been a topic of considerable interest for at least thirty years (Allen & Graden, 1995; Deno, 1985; Fuchs & Deno, 1991; Graden, 1989; Graden, Casey, & Chistenson, 1985; Ikeda, Tilly, Stumme, Vollmer, & Allison, 1996; Macmann, Barnett, Allen, Bramlet, Hall, & Ehrhardt, 1996; Pugach & Johnson, 1989; Reschly & Ysseldyke, 1995; Zins & Erchul, 1995). Numerous approaches have been proposed and tested across various different domains of school functioning, typically with an emphasis on improving student performance in some capacity. Formalized problem-solving processes have been applied to problem behavior (O'Neill, Horner, Albin, Storey, & Sprague, 1996; Tilly et al., 1998), academic performance (Deno, 1985), and prevention (Good & Kaminski, 1996). The What Works Clearinghouse listed data-based decision making as an evidence-based practice within the Response to Intervention Framework (Gersten et al., 2008), and research on decision-making within schools has provided a great deal of insight into the parameters of effective and efficient problem-solving.

Bergan and Tombari (1976) conducted a study of the relationship between proficiency with consultative problem solving and several outcomes. In their study, they looked at problem identification, solution implementation, and problem resolution. They found that proficiency with consultative problem solving was associated with a 43% incidence of problem identification in those cases studied, 31% incidence of solution implementation, and 30% incidence of problem resolution. It is unclear from their reporting of findings the degree to which problem resolution was associated with the other two components. The essential logic model of all problem-solving practices would seem to stipulate that the effect of problem solving on problem resolution should only be observed when both: (a) a problem has been identified, and (b) a solution has been implemented. If this relationship does not bear out in the data, it must appear that some other aspect of training is exerting the observed effect on problem resolution (i.e., greater facility with interpersonal skills as taught in the training). Their study did not report the degree to which these variables co-varied. Additional mediational analysis would be required to clarify this issue.

Kratochwill, Elliott, & Busse (1995) conducted a study with 17 graduate student-consultants examining the effect of consultation training on a variety of problem-solving consultation skills. Key take-aways from the study were that training was highly effective in producing adherence to the model by the graduate-student consultants, and the mean effect size of interventions devised through consultation was .95. Of additional note, just over one-third of cases consulted had no positive student outcome to report. These findings then seem to indicate that when applied with fidelity, this model of problem-solving has the potential to have a very meaningful impact for a simple majority of

children for whom it is used to design interventions. However, these findings also indicate that this model of consultative problem-solving does not appear to be effective for producing positive results with all students, thus impacting the degree to which it can be expected to scale within educational settings. Generalizable problem-solving strategies suitable for wide-range scaling within education should be expected to produce some kind of positive effect for all children for whom they are used.

Telzrow, McNamara, & Hollinger (2000) conducted a study examining the effect of problem-solving fidelity (adherence to the steps advocated in formalized problem-solving) on student outcomes. The authors found that demonstration of problem-solving components was positively associated with student outcomes as rated by researchers who observed multidisciplinary team meetings focused on student problems. Findings indicated that those elements of the problem-solving process associated with data use and decision-making regarding interventions on the basis of data were comparable to observed demonstrations of student progress during meetings in terms of predicting positive student outcomes. While promising, only 8% of the variance in subjectively rated student outcomes was accounted for by the strongest elements of problem-solving fidelity. This indicates that additional work needs to be done in this area with objective measures of student outcomes and rigorous criteria for problem-solving implementation.

Data-based decision making is a critical component of effective academic support delivery in general (Hattie, 2008), and a core feature of Response to Intervention (RTI; Fuchs & Fuchs, 2006). When schools consistently utilize data to inform decision making related to student supports, students achieve higher levels of proficiency at higher rates (Fuchs & Fuchs, 2006; Hattie, 2008). Several models have been applied to decision

making related to school-based supports specifically (Deno, 1985; Fuchs & Fuchs, 2006; Good & Kaminski, 1996; Todd et al., 2011). Several models have been developed with a focus specifically on literacy instruction and supports (Deno, 1985; Fuchs & Fuchs, 2006; Good & Kaminski, 1996). Of those focused on literacy, some have focused on the efficient remediation of deficits (Fuchs & Fuchs, 2006), and others more on the prevention or mitigation of deficits (Good & Kaminski, 1996), but all have emphasized the need for efficient data systems for tracking and evaluating students' progress mid-year (Deno, 1985; Fuchs & Fuchs, 2006; Good & Kaminski, 1996).

Fuchs and Fuchs (1989) conducted a study examining the effectiveness of decision making training using a consultant model. While the authors note a high level of integrity implementing problem identification and intervention design components of the problem-solving process, they observed a much lower rate of processes associated with data usage (24-39%). Notable in this study is the lack of follow-up support available to consultants and the 1:1 support model (one consultant, one teacher) used in the study. While one could conceptualize a consultant as providing coaching support, the model applied in the study did not meet this criterion regarding generalized decision making because emphasis was placed on supporting teachers in their work with a single student, not on supporting their problem-solving across students or over time (both necessary for coaching a generalized problem-solving repertoire). Further, the lack of a team in the decision-making process is a potentially critical difference in terms of both interpreting the lack of data use observed and in terms of generalizing the positive portions of their findings to current school settings. Modern school-based decision-making is predominantly team-mediated.

Another major issue which has been raised in the literature on school-based decision-making is a serious lack of both quality and fidelity in data collection activities (Flugum & Reschly, 1994; Fuchs & Fuchs, 1989; Telzrow, McNamara, & Hollinger, 2000). Prior to training, and in some cases even after training, data collection is limited thus attenuating the degree to which decisions can be defensibly based upon such information. One major issue in the drive to improve decision making in schools must then be to facilitate teams' and teachers' efficient collection of high quality data. To achieve this goal, teams must at a minimum be trained and supported to: (a) generate observable problem statements for which high quality data can be collected, (b) plan data collection activities explicitly, (c) efficiently design such data collection procedures as will facilitate formative and summative evaluations regarding interventions targeting a specified problem, and (d) incorporate systematic review of such data into their formalized decision-making process for both formative and summative decision making at regular intervals.

Glover (2017) defines the DDIC in terms of three domains: (a) formalized decision making, (b) a standardized coaching process guiding all other aspects of implementation, and (c) specific domains of coach-delivered teacher support. The author explicates a four component decision making process which is very much consistent with other models that have been identified in both the broader problem-solving literature and within the more narrowly focused literature pertaining to problem-solving within educational contexts. The four components of the DDIC problem-solving process are: (a) problem identification, (b) problem analysis, (c) action plan implementation, and (d) evaluation of goal attainment. Consistent with the behavioral consultation model (Bergan,

1977; Bergan & Kratochwill, 1990) upon which it is based, the DDIC emphasizes the putative mediating effects of teacher perceptions on teacher practice and student outcomes. The three features of coaching identified as critical within the DDIC: (a) emphasis on the learning environment (interactionally situated between teacher and learner), (b) coach delivered modeling, practice and performance feedback for teachers on targeted skills and strategies, and (c) a formalized decision-making process. Finally, four aspects of teachers' instructional behavior are identified as being of critical importance within the model: (a) academic screening, (b) identification of students' *specific* skill needs, (c) homogeneous grouping based on *specific* skill needs, and (d) progress monitoring.

Across contexts and foci, the steps of effective decision making have remained quite consistent over time. Despite these consistencies and considerable investment in research on decision making in general, empirical demonstrations of efficiently structured, readily-scaled decision-making processes for academic behavior have remained elusive. While it is clear that the decision rules applied to academic problem solving have significant potential for remediating and preventing student reading challenges (Coyne et al., 2002; Deno, 1985; Fuchs & Fuchs, 2006; Glover, 2017; Good & Kaminski, 1996), application of these rubrics within a scalable structure awaits verification (Balu et al., 2015; Foorman et al., 2016; Gersten et al., 2008; NAEP, 2015; Spectrum K-12, 2010). One approach to verifying an effective approach to this issue is to leverage those practices that have been successfully applied to other content areas within schools.

While much of the early work on problem-solving within schools emphasized this type of behavioral consultation model, approaches over the last twenty years have demonstrated increasing appreciation for team-based models of problem-solving (Allen & Graden, 1995; Chalfant, Pysh, & Moultrie, 1979; Macman et al., 1996; Pugach & Johnson, 1989; Todd et al., 2011). Two excellent examples of where team-based problem-solving has been used and tested extensively are RTI and PBIS. Within PBIS, teams meet regularly with the purpose of reviewing data to identify and troubleshoot problems associated with students' social behavior. To facilitate efficient problem-solving during such meetings, a team of researchers developed the Team-Initiated Problem-Solving model (TIPS; Newton, Horner, Algozzine, Todd, & Algozzine, 2009). The TIPS model was designed for application within the PBIS framework, emphasizing the elements of efficient decision-making with teams who had already received training in Tier I PBIS. The TIPS approach to problem-solving provides a well-defined structure for integrating a teams' content expertise with their utilization of relevant data systems. TIPS provides a structure for using these data systems to improve implementation of evidence-based systems and practices, and builds from a long history of DBDM efforts in education and psychology.

Team-Initiated Problem-Solving. Using this approach, the research examining Team-initiated Problem Solving (TIPS) as applied to social behavior in schools is particularly promising. TIPS has all the hallmarks of effective problem solving embedded within its six step Problem-Solving process as well as the essential preconditions for effective team-based work in its Foundations subscale. Research on TIPS indicates that it

is technically adequate in terms of both reliability of its subscales, and in terms of its content validity (Algozzine, Newton, Horner, Todd, & Algozzine, 2012).

Training in TIPS for PBIS teams has been standardized using the TIPS training manual, manualized slides and activities, and post-workshop coaching with in-district coaches who themselves are trained in supporting TIPS implementation. Each of these components is presumed critical to the successful implementation of TIPS as observed in prior research (Newton et al., 2012). While there is an extensive research base supporting each of these components, the addition of coaching to the process of decision making training is likely the most critical to the unusual level of success observed with TIPS implementation following training with coaching.

Early research on TIPS included a small pilot study (Todd et al., 2011), and development of the first edition of DORA (Newton et al., 2009). In their pilot study of the TIPS model, Todd and colleagues worked with four Title I elementary schools in Oregon that were already implementing SWPBIS and using the School-wide Information System (SWIS; Horner et al., 2008; May et al., 2003). Extensive observational data were collected using DORA to document the degree to which each team exhibited each aspect of the problem-solving process emphasized in TIPS. The team used a multiple baseline across teams design, collecting a minimum of six baseline data points prior to intervention with TIPS training, and two data points after intervention. The researchers concluded that TIPS training exhibited a functional relation with foundational aspects of meetings and implementation of the TIPS problem-solving process, as indicated by visual inspection of the data. This then left the question of the degree to which these results could be generalized more broadly.

In a larger-scale follow-up study, this same research team used a block-randomized waitlist-controlled design. In this follow-up, Newton and colleagues (2012) demonstrated that TIPS training resulted in improved implementation of research-based problem-solving processes in a larger sample, using randomization as a control for potential confounds. Their sample included PBIS teams from 34 schools in Oregon and North Carolina, each of which had been implementing PBIS for at least one year prior to participating. Implementation of research-based problem-solving processes lacked many essential features at baseline, and lacked adequate implementation of most features. Following training, teams improved their performance as measured on the Decision Observation, Recording, and Assessment tool (DORA) by approximately 1.7-2.0 times the number of rubric points. Rubric points on the DORA/DORA-II are tied to the execution of specific steps in the problem-solving process (though some items are dependent on others, creating a small amount of local dependence). The findings from this study indicate a clear functional relationship between TIPS training and implementation of a research-based problem-solving process as applied to social behavior. While this study convincingly demonstrated that manualized TIPS training exhibits a functional relation with problem-solving process integrity across schools of varying features it did not address the impact of TIPS on student outcomes.

In the most recently published RCT on TIPS, Horner and colleagues (2018) followed up on the 2012 study with a sample of 38 school teams from schools in North Carolina (n=20) and Oregon (n=18). This team of researchers documented the successful implementation of the TIPS model following training for both the immediate training group and the waitlist group. Problem-solving scores as measured on DORA-II were

statistically significantly higher for the immediate group relative to the waitlist group for observations directly after TIPS training for the immediate group ($p = .005$). Further, differences at this observation demonstrated a large size of effect ($ES = .96$). Following training of the waitlisted teams, both the immediate group and the waitlist group performed with similarly high problem-solving scores ($M = .82, .79$, respectively). Sixteen of the nineteen schools in the immediate group also had fewer office discipline referrals following training as compared with only 10 of 19 in the delay group at this time. Further, the immediate group demonstrated a statistically significantly lower rate of out of school suspensions during the final observation period, as compared with the delay group. The research base supporting implementation of TIPS to decision making related to social behavior supports is strong. However, it remains to be seen how well: (a) TIPS training improves implementation of research-based problem-solving processes applied to academic behavior, and (b) to what degree implementation of such a process will impact student academic outcomes.

From this perspective, there are at least seven elements of data-based decision-making relevant to applications of academic achievement within public schools (i.e., six-step problem-solving process, logistic foundations). Of critical importance to an analysis of data-based decision making within schools is the practice of teaming. Teaming is the practice of bringing together a group of professionals from within a given setting for the purpose of making decisions. Teams within public schools have several distinct advantages. First, they allow a school to leverage its full breadth of expertise by bringing together individuals with different skill sets (i.e., reading specialists, special educators, etc.) or domains of expertise (i.e., school psychologists, behavior specialists, etc.).

Further, teams facilitate problem appraisal by multiple individuals, a process that reduces the likelihood of missing critical information and simultaneously increases the likelihood that irrelevant information will be appropriately ignored. All three of these advantages are enhanced by high levels of functioning across two aspects of team based problem solving: efficient logistics, and effective, systematic problem solving.

Logistics of team meetings include elements like starting on time, ending on time, consistency of meeting schedule and attendance, the method of sharing critical information, and disseminating an agenda to the group to guide discussion. These elements of meetings must be efficient to facilitate sufficient time for problem solving each issue raised during a given meeting. Additionally, efficient logistics reduces wasted time during meetings by ensuring that critical information is shared effectively, team members are typically present, and discussion is guided effectively via a shared agenda.

Data based decision making involves five or six steps, depending upon how finely the process is parsed. These steps are: (a) problem identification, (b) solution proposal, (c) solution implementation, (d) solution monitoring: implementation, (e) solution monitoring: impact, and (f) summative evaluation. For efficient decision making, it is critical that the first step is done with precision on the basis of data. Data for informing problem identification are best when they are observable and replicable in their methods. The precision required for problem identification must be sufficient to differentiate this problem from other similar problems in terms of where it occurs, when it occurs, what the core issue is (what makes it a problem), why it is occurring (a functional hypothesis), and how it may be solved (solution proposal should flow logically from problem identification on the basis of hypothesized function).

The Present Study

Response to Intervention is widely implemented with partial fidelity, but a significant challenge remains in providing school teams with the training they need to effectively implement its data-based decision-making components (Foorman et al., 2016; Gersten et al., 2008; Hoffman, Jenkins, & Dunlap, 2009; Sharp et al., 2016). Effective decision rubrics have been identified for substantially improving student outcomes within reading (Coyne et al., 2004; Good & Kaminski, 1996; Glover, 2016; Harn et al., 2011). However, one persistent challenge has been training teachers in the sustained implementation of these rubrics for instructional decision making (Hoffman, Jenkins, & Dunlap, 2009). Effective implementation of a codified process of data-based decision-making has been successfully implemented within the related domain of PBIS. It is likely on the basis of the similarities of these systems that such a process could be successfully applied to the academic decision-making within RtI as well.

Toward this end, the present study sought to address the following research questions:

- 1) Is there a functional relation between exposure to Academic TIPS training and increase in DTs' implementation of the TIPS model: (a) meeting foundations, (b) development of precision problem statements, and (c) design of technically adequate academic support plans as measured by DORA?
- 2) To what degree is implementation of literacy supports guided by the TIPS Model associated with changes in student literacy (as assessed using Easy CBM)?

- 3) To what degree is the TIPS approach perceived as acceptable to members of DTs as measured by the Adapted Self-Assessment of Contextual Fit and the Primary Intervention Rating Scale (Horner, Salentine, & Albin, 2003; Lane, Robertson, & Wehby, 2002)?

CHAPTER III

METHOD

Participants

Primary participants for this study were the five members of the core data team (DT) from one elementary school (K-5) in Western Oregon. The DT was assembled by the principal with the purpose of monitoring student academic progress, and designing literacy interventions to improve student success. The DT was composed of the reading specialist for the school, two administrators, and two special educators. Members of the core DT met nearly weekly, with the focus of each meeting (and additional teaching staff) rotating to address students in a selected grade level (grades K, 1, 2, 3&4, 5). This schedule resulted in the DT meeting with the members of each of five grade level teacher teams (GLTT) approximately once every five weeks to review literacy data and problem solve solutions. Each grade had its own GLTT with the exception of grades 3 and 4, which were combined due to the presence of a single 3/4 blended classroom. Each GLTT was comprised of 3-5 general education classroom teachers. Each observed literacy meeting included the four core DT members and the supplemental GLTT members. A total of 26 teachers and related staff participated in the study (4 core DT members, an additional administrator, and 21 GLTT members).

The data team members for the study were recruited from a convenience sample of schools currently using PBIS (Horner, Sugai & Anderson, 2010) who expressed interest in participating when solicited through a written and/or email invitation. Participating teachers were contacted for participation in connection with their typical

teaming and professional development activities (at regularly scheduled team meetings or in-service trainings). The selected DT reviewed informed consent information for the proposed study and each member consented to participate during regularly scheduled meeting/professional development times. If a DT or GLTT member did not consent to participate, then their contributions during team meetings was not be recorded or used during any data collection activities. All DT and all but one GLTT member chose to participate. The DT facilitator was instructed to re-iterate information contributed by this member and in the event that the team had any guests during the course of the study.

Teachers at the participating school were predominantly White females. Total enrollment at the participating school was 485 students (K-5). The student body at the participating school was 75% White, 20% Hispanic, 2% Multi-racial, 1% Native American/Alaskan Native, 1% Native Hawaiian/Pacific Islander, and 1% Asian. A majority of students came from families meeting the federal criterion for economically disadvantaged as indicated by qualification for free or reduced price lunch (70%). Significant segments of the student body were mobile (14%), had limited English proficiency (13%), or required an individualized educational plan (15%).

Student level literacy data were collected from permanent products. Specifically, EasyCBM screening and progress monitoring data, which were routinely collected three times per academic year (fall, winter, and spring) for all students and every one-to-two weeks for students with increased risk status. Passive consent was obtained for all eligible students (i.e., K-5).

Setting

The participating school had been implementing multi-tiered systems of support for both behavior (PBIS) and literacy (RtI) for at least the last three years. Fidelity of implementation for both behavioral support systems and academic support systems has been a challenge and recognized area for growth during that time. Data from the Tiered Fidelity Inventory (Algozzine et al., 2014; McIntosh et al., 2017) indicated that they were partially implementing PBIS (i.e. not to criterion). Criterion for implementation with fidelity on the TFI is .70. On the Tier I portion of the Tiered Fidelity Inventory the school achieved a score of .50. The school's score for Tier II was .27, and .47 for Tier III. Each of these scores indicated a need for improved fidelity of implementation. Of particular interest were scores for the subscales related to data usage. The highest rating for items corresponding to data usage and decision making was a 1 out of 2, indicating that a data system was in use and data were reviewed for decision making, but that the data system was not well-understood and decision-making occurred at low frequency. This is interpreted as indicating that the basic requirements for data-based decision making were in-place but additional staff training was needed to render this practice feasible and effective in this setting. No data were available for fidelity of RtI implementation.

Most team meetings and study related activities were carried out in teachers' classrooms, approximately 20' X 20', each with an interior door, an exterior door, several windows, 2 whiteboards, approximately 25 student desks, a document camera, a teacher desk, two small group work tables in the back, an additional small room contained within the larger classroom often used for student breaks, and other classroom related materials. During team meeting observations, observers were positioned behind all team members,

in view of any information projected during the meeting. Training activities occurred in similar rooms.

Problem behavior at the participating school during the 2017-18 school year was slightly higher than the national median, in terms of office discipline referrals per 100 students per school day (.3 versus .2) as assessed using the School Wide Information System (SWIS; (May et al., 2003). A majority of students did not meet performance standards in reading (56.5%) as assessed using the Smarter Balanced assessment for end of year testing (Consortium, 2016). A similar proportion did not meet performance standards in mathematics as well (62.4%) as assessed using the Smarter Balanced assessment for end of year testing (Consortium, 2016). Similar results are obtained when using EasyCBM benchmarking assessments as the criterion of evaluation (Alonzo, Tindal, Ulmer, & Glasgow, 2006).

When controlling for entry scores, median growth was observed for students at the participating school on average (52nd percentile relative to other schools in Oregon) in reading, but not mathematics (41st percentile). This means that relative to students at other schools within the state of Oregon who had similar reading scores at the beginning of the school year, student growth was at the 52nd percentile. This indicates that their students' response to intervention, given their initial level of performance, was typical of other students within the state of Oregon. Growth for non-majority students was lower for most categories. Students with special learning needs demonstrated growth at the 26th percentile relative to other students who entered with similar scores within Oregon. This means that students at this school who received special education services demonstrated less growth than their peers with similar entry skills at 74% of Oregon schools. Students

from underserved racial backgrounds demonstrated growth at the 42nd percentile, and economically disadvantaged students grew consistent with the 41st percentile. English learners grew at very near median performance (47th percentile). In math, the same methods resulted in scores of 30th percentile growth for economically disadvantaged students, 37th percentile for English learners, 20th percentile for students with special learning needs, and 38th percentile for students of underserved racial backgrounds. Overall, these data indicate that students from non-majority backgrounds improved their reading skills at slightly to moderately lower rates than their similarly performing peers at other Oregon schools.

Measures

Decision Observation, Recording and Analysis (DORA-II). The Decision Observation, Recording and Analysis (DORA-II) tool measures core elements of the TIPS problem solving model. DORA-II has two primary sections: (a) Foundations, and (b) Problem Solving (with two subscales: Decision-making and Implementation). Initial demonstration of the technical adequacy of DORA-II was conducted with schoolwide behavior teams (Todd et al., 2011). Technical adequacy of DORA/DORA-II for measurement of data-based decision making quality in the context of schools teams has been demonstrated across two waitlist-randomized control trials and numerous smaller scale studies (Algozzine, Newton, Horner, Todd, & Algozzine, 2012; Horner et al., 2018b; Horner et al., 2009; Newton et al., 2011; Newton et al., 2012a; Todd et al., 2011). DORA-II has demonstrated validity and reliability for measuring team problem solving, with inter-observer agreement of 97% for Foundations and 90% for Problem solving (Horner et al., 2018a). DORA-II scores for the proposed study were decomposed and

summarized into three subscales: (1) Foundations, (2) Decision Making, and (3) Comprehensive Plan with Implementation. These will each be comprised of specific DORA items relevant to critical aspects of the training as organized for the proposed study.

Foundations. The Decision Observation, Recording and Analysis II tool (DORA II) has a section composed of 10 items related to the presence of essential foundational features of team meetings. Items focused on meeting foundations included: initiating the meeting (starting on time, use of a publically shared agenda, access to previous meeting minutes, and attendance at the start of the meeting), roles during the meeting (facilitator, minute keeper, data analyst), and ending the meeting (ending on time, attendance at end of meeting, scheduling the next meeting). A subscale score for DORA-II Foundations is calculated as a percentage of 10 items correct, ranging from 0% to 100% (Algozzine et al., 2018).

Decision Making. The DORA-II tool has a section composed of 20 items that assesses implementation of data-based decision making consistent with the TIPS model. A portion of this section, Problem Solving (11 items), was used for measuring the degree to which teams demonstrate the critical features of data-based decision making by implementing the TIPS problem solving model. Items measured on the Problem Solving portion of the DORA-II include: 1) Problem Defined with Precision (Who, What, Why, Where, When), 2) Problem Category (Social, Academic), 3) Problem Features (Group/Individual, New/Old), 4) Quantitative Data Use (Social Behavior, Academic Behavior), 5) Goal Identification (Change criterion, Timeline).

The remaining nine items measure the degree to which the team demonstrates the critical features of implementation planning and evaluation. These items include: 1) Solution Identification, 2) Solution Implementation (Person responsible, Timeline, Solution ID, Timeline with Goal, Evaluator ID), 3) Implementation Integrity (None, Partial, With Integrity, Stopped, Unknown), 4) Problem Status (Worse, No Change, Improvement below goal, Goal met, Unclear, Unknown), and 5) Summative Evaluation Decision (NA, Yes, No; Retaining, Revising, or Terminating a solution/goal/problem/combo). This tool allows for item level summaries, or summaries by problem solving feature addressed by multiple items (i.e., data-based problem identification encompasses items 1-5). DORA-II score is calculated as the percentage of these nine items completed by the team during problem solving meetings.

For the purposes of the current study, each element of a precision problem statement (5 for social behavior, 3 for academic behavior), the problem category and type of quantitative data used (3 items), the problem features (3 items; New/Old problem, affecting an individual, group, or both), and a goal statement including what change is to occur by what date to resolve the problem under consideration were used for calculation of the Problem Solving score. The Problem Solving score was calculated as the percent of these items addressed during the DT meeting.

Implementation Planning and Evaluation. The implementation planning and evaluation section is comprised of the final 9 items of this section of DORA-II. These include seven items relevant to existing problems, and 5 items relevant to new problems. Relevant to new problems are the items associated with the Solution Implementation Plan section of DORA-II, which includes 5 items: (1) Person responsible for implementing the

plan, (2) Implementation Timeline, (3) What Treatment Integrity will be collected, (4) When Treatment Integrity will be collected, and (5) Who will collect Treatment Integrity data. For existing problems, these issues have already been addressed at previous meetings and are thus no longer relevant. Monitoring treatment integrity data and the impact on students is very much relevant for existing problems however, and these are the focus of this section for existing problems. On DORA-II this includes sections on Solution Implementation Integrity, Status of Problem, Comparison to Goal, and Summative Evaluation Decision. Solution Implementation Integrity includes 5 response options for a single item relevant to existing problems, each requiring that the team review data and evaluate the fidelity of implementation for the plan in question. Status of Problem includes whether the status has been reported on, and whether data were presented relevant to that status (2 items). Comparison to Goal is a single yes/no item indicating whether the team has compared progress against the stated goal explicitly. Finally, Summative Evaluation includes a yes/no indicator of whether a decision has been made and three sub-items indicating what the features of that decision are. Each plan is evaluated using available data and teams are tasked to determine whether the solution, goal, or problem statement should be adjusted. For each of these items, teams can choose to: (1) retain the solution/goal/problem statement, (2) revise the solution/goal/problem statement, or (3) terminate the solution/goal/problem statement. When progress is being made effectively, but not yet to the goal criterion, teams should typically persist with the existing set of parameters. When progress has met the goal, termination of the plan should be considered (though not guaranteed, as the plan may be needed for continued success). When the plan has produced too little or no progress, some or all aspects should

be revised. For the proposed study, the total number of items present for all new problems discussed during a given meeting were summed with those for old problems, divided by the sum of items possible ($[7 * \text{number of old problems}] + [5 * \text{number of new problems}]$), and multiplied by 100 to get the percent correct for this subscale.

Student Behavior. Curriculum-based measurement in reading (CBM-R) was used to assess student reading achievement growth over the year as a function of DT decision making. CBM-R includes measures of fluency for: oral passage reading, word reading, nonsense words, phonemic segmentation, letter names, letter sounds, and segmenting phonemes. Additionally, reading comprehension as measured using multiple choice assessments were used for older students. Major examples of CBM-R packages utilized nationally include: Dynamic Indicators of Basic Early Literacy Skills (DIBELS), EasyCBM, and AimsWEB. CBM-R is suitable and commonly used for both universal screening and progress monitoring. Each of these measurement suites has been validated as a set of measures which are valid and reliable in assessing student reading progress over time (Alonzo et al., 2006; Good III & Kaminski, 1996; Shinn & Shinn, 2002). EasyCBM is the CBM-R suite used in the participating district and includes an estimate of overall reading risk based upon all administered measures. This estimate of overall reading risk was used in addition to the individual metrics administered.

Social Validity. A teacher survey was used with the five DT members to measure the acceptability of Academic TIPS. The Adapted Self-Assessment of Contextual Fit (Horner, Salentine, & Albin, 2003) is a 16 item survey using 1-6 Likert scale ratings to assess the degree to which a given practice fits the local school context within which it is being applied. The ASACF was used as an overall indicator of social validity. This

measure has served as an acceptable indicator of social validity in past work within the domain of educational practices (Monzalve, 2016). As an additional indicator of social validity, the Primary Intervention Rating Scale (PIRS) was used (Lane, Robertson, & Wehby, 2002). The PIRS includes 17 Likert items (1-6 scale) and four open-ended questions. It has been validated for the evaluation of acceptability of systems and procedures being implemented by entire schools.

Academic TIPS (AcTIPS) Training

Training included three two-part sessions. The first session was 60 min in length, the second and third sessions were each 90 min in length. The target audience of the trainings were the five members of the DT. Some additional school staff members attended each training at the request of the building principal to facilitate implementation of TIPS for the school's PBIS teams. For each session, the training portion was half to two-thirds of the allotted time and was immediately followed by an application portion using each team's own data for a simulated team meeting focused on the portion of the TIPS model covered during the preceded training portion of the session. Each training session covered a different portion of the content on the TIPS model and its application to literacy, numeracy, and social behavior problems. The training portion of sessions was roughly 60% didactic presentation of content and 40% practice with performance feedback. The training sessions are based off of prior work in this area (Newton, Todd, Algozzine, Horner, & Algozzine, 2009), with modifications to the content designed to emphasize academic performance (literacy) as this constitutes the majority of problems typically addressed during GLTTs. Additional modifications were made to emphasize the areas of specific need for typical GLTT members. The most significant of these changes

was the expansion of the segment on data analysis and summarization due to the relatively sparse training most general education teachers receive on this topic during either pre- or in- service trainings (Appendices A & B). Specific training content for each of the three sessions is provided below.

Session 1. The first training session covered TIPS Foundations, and focused specifically on Team Roles. Foundations of effective team meetings includes the following factors: adequacy of team membership, identification of team roles, use of public agenda, access to relevant quantitative data, regular scheduling and time-keeping (advanced scheduling, at least once every six weeks, with regular start and end times), identification and agreement of decision rules prior to meetings, and identification and agreement of malleable factors (what interventions and changes the team can make to reading supports provided to students). Adequacy of membership is observed when at least three grade level teachers are regular attenders of the meeting in question. Identification of roles is observed when, prior to the meeting start time, the three critical roles for data-based decision making are identified: facilitator, data analyst, and minute taker. Execution of roles is observed when an individual member of the team serves each role independently. The facilitator prepares for the meeting by reviewing the prior meetings minutes and setting the agenda for the current meeting at least one day before the team meets. Inclusion of old problem review and new problem solving segments are critical elements of every agenda and serve as indicators of facilitator role execution. Minute takers record the results of old problem review and new problem solving, including analysis of progress to date, revisions to existing plans, new precisely defined problems, goals for student/group achievement, and solutions with plans for

implementation. Execution of this role requires entry of the critical information from the TIPS problem solving process into a form for tracking and reference. Execution of the data analyst role involves reviewing relevant data, as indicated on the previous meeting minutes form or related to screening for new problems, and prepares for displaying data during the meeting one to two days prior the scheduled time. During the meeting, the data analyst graphically presents (or identifies for the facilitator) data relevant to past or current problem solving for the team to review. The data analyst then provides a summary statement for each level presented in the graphed data (i.e., small group and student, letter naming fluency and word reading fluency, etc.).

Access to relevant quantitative data is observed when graphed data are displayed or shared via printed copy during the meeting, without the need for any team member to leave the room to gather or gain access to this information once the meeting has started. Regular scheduling is observed when the next scheduled meeting is displayed in meeting minutes or known to all meeting members as indicated by verbal query. Regular time keeping is observed when the team begins within ten minutes of the scheduled start time, and ends within ten minutes of the scheduled end time or agrees by unanimous query and consent prior to the end of the meeting to extend the end time to resolve a particular task. Identification of decision rules involves the team's collective agreement on what constitutes criterion/expected performance across domains the team will evaluate during the meeting, and what type of decisions are made when a student or group of students are not demonstrating criterion performance as indicated by data the team regularly reviews. Similarly, identification of malleable factors involves the agreement by all members of the team, prior to the meeting start time, regarding what aspects of the instructional

environment are subject to change based upon the team's collective judgment. These latter two elements are often informed or dictated by administration at the school or district level. Whether determined by administration or exclusively by members of the team, a written summary of both decision rules and malleable factors over which the team can exercise control should be present either in paper or electronic format during each team meeting.

During this session, each GLTT coordinated with the other staff members on their DT to determine which members of each DT iteration would serve each role. Specialized breakout training was provided to participants by role, with those not identified for a specific role spread across the three role-based groups to receive training as backups. Training for each role provided each participant with the critical documents and systems related to their new role, provide modeling, practice opportunities, and performance feedback on their execution of their new role in a scaffolded setting.

Data analyst training emphasized gathering and displaying data that meet two conditions: enable decision making in line with the group's pre-identified decision rules or that are relevant to following up on the problem solving done at the previous meeting (fidelity of intervention delivery, diagnostic assessment as needed), and are capable of informing all relevant aspects of a precision problem statement (informing the who, what, and why for academic problems, etc.). Data used for examples were consistent with the data systems used at the participating school. These systems used in the training included: EasyCBM and School-wide Information System (SWIS). The data analyst was further instructed to provide a summary statement for each graph presented during team meetings that included: labeling the x-axis variable and scale, labeling the y-axis variable

and scale, describing the level of the group, describing the trend of the group, identifying any individual students who differ in their performance from that exhibited by the group, and describing the level and trend of performance for such students.

Facilitator training emphasized setting and moving the team through the agenda each meeting. This involves estimating the time required to complete each stage of the meeting in advance of each meeting, directing and re-directing the conversation to malleable factors related to the topic at hand during topical discussions, and building consensus to facilitate action. Facilitators were presented with example documents that were pre-filled with estimated times for typical meetings, and then given blank materials with a set of meeting objectives and coached through the process of setting an appropriate agenda. Finally, facilitators were given an opportunity to get feedback on a proposed agenda for an upcoming team meeting using their teams' current information (data, priorities, etc.). The facilitator was also given a job aide prompting the relevant questions associated with application of the TIPS process to academic content.

Minute keeper training focused on targeted note-taking. Targeted note-taking during TIPS meetings provided information critical to decision making using the TIPS model. In particular, most information regarding non-malleable factors is omitted and information related to specific aspects of a precision problem statement, potential solutions, or plans for implementing solutions is emphasized. The TIPS GLTT meeting minutes form readily guides this process and was used to guide instruction and practice activities during this breakout. Meeting minutes taken proficiently enable a third-party observer, after the meeting and without attending, to identify: (a) precisely what type of problem solving occurred, (b) for which students/student groups, (c) what problems were

identified precisely, (d) what solutions were selected to address these problems, (e) who will implement the solutions, (f) when the solution were implemented, (g) when the problem performance is expected to be resolved, (h) who will monitor fidelity of implementation, and (i) how fidelity was monitored.

Session 2. The second training session began with a brief review of Session 1 content. New content emphasized use of academic data, but was balanced with examples across literacy (~40%) numeracy (~30%), and social behavior (~30%) data. This session included information on identifying problems with precision and identifying goals for support. Precise problem statements are critical to efficient data-based decision making because they reliably lead to interventions which are matched to students' needs (TIPS citation). Precise problem identification requires basic data analysis skills for all members, 6 components of precise problem statements, and goal-setting. Basic data analysis requires display of appropriate data and summarization and interpretation of essential features. The six components of precise problem statements are: (a) Who is the problem affecting, (b) What is the problem (skill of deficit, degree of deficit), (c) Why is the problem continuing to occur (hypothesized mechanism or behavioral function), (d) When the problem is most likely and least likely to occur (for problems with a social component), (e) Where the problem is most likely and least likely to occur (for problems with a social component), and (f) how consistently the problem is occurring (i.e., how often social problems occur or what proportion of opportunities result in errors for academic problems). Regarding the "Why" for a particular problem, it is important to note the similarities and differences between social and academic behavior. Within social behavior (the domain within which TIPS was first validated), Why typically emphasizes

the *function* of the behavior (Newton et al., 2009). Within the context of academic behavior, the student's proficiency with prerequisite skills is typically emphasized. For every behavior (desired or problem/error) a student emits during the course of their school career, whether it is of primarily academic or social relevance, there are two forces that drive that student's proficiency emitting the *desired* behavior: motivation and prerequisite skills. Motivation refers to the degree to which the student is motivated to engage in effortful responding to obtain the presumptive consequence, or sometimes the amount of effort a student is willing to expend to obtain that consequence. This is most salient when discussing the function of socially-relevant problem behavior. The nature of a functional relation in the context of problem behavior is such that the functionally related consequence *maintains/supports* the occurrence of the problem behavior or alternative behavior. This observation is a clear indication that the student is motivated to work for that consequence.

Prerequisite skills include all responses which the student must emit in the course of performing the desired behavior, *or* in order to learn how to emit the desired behavior. Traditionally, schools have emphasized the role of motivation in the context of socially relevant behavior (problem behavior in particular), and prerequisite skills in the context of academically relevant behavior. For any given student, a problem of social or academic performance may derive from issues associated with either motivation or proficiency with prerequisite skills. Both are mutually supportive in many cases as illustrated by the example of a student with mildly impaired reading skills and low classroom motivation. This student may find reading effortful while being fully capable of criterion performance. The contingencies associated with engaging that performance

may not be satisfying to the student (i.e., a token delivered silently for a student who desperately wants more teacher attention). This situation can result in two related problems: (a) the student engages in sub-criterion reading performance, or (b) the student engages in escape-maintained problem behavior to avoid the reading task. The sequelae of either of these immediate responses is that the student will acquire greater difficulty with reading at criterion over time, likely to the point of developing a true reading deficit. Alternately, a student may lack the social skills to interact with peers in a manner which garners positive attention. As a result, the student may engage in inappropriate attention-seeking behavior such as classroom disruption. In this instance, the student is motivated to obtain peer attention and lacks the skills necessary to obtain such attention in an appropriate manner. Addressing such a student's problem behavior will necessarily include skill building.

An imprecise problem statement includes superficial information about a problem, "Johnny is having trouble with reading..." instead of a precise statement like, "Johnny (Who) is struggling with word reading (What). He also has consistently scored low on phonemic segmentation and letter sounds (how often). His scores for the most recent period place him in the 15th percentile for word reading (to what degree), the 10th percentile for phonemic segmentation, and the 12th percentile for letter sounds. We think Johnny is struggling with word reading because he doesn't know all of his letter sounds fluently and he still struggles to break words into parts (Why)." The former indicates that Johnny needs additional help learning to read. The latter indicates that Johnny needs instruction in sound-symbol correspondence focusing on letter sounds and phonemic awareness focusing on segmenting. The first conclusion is entirely unhelpful to planning

specific intervention supports, even with information about Johnny's word reading fluency score it would not be much better. The second example indicates a clear, specific need, and suggests a plan for remediating Johnny's deficit.

Goal-setting involves 3 steps, identifying the performance criterion indicating that the problem has been remediated, the timeline for fully remediating the problem, and a specific statement of what change in the target behavior/skill will occur by what date (if full remediation is not targeted in the present meeting). In the context of reading this may be, "Johnny will improve his letter sound fluency to meet criterion, and his phonemic segmentation and word reading fluency to performance above the 20th percentile in 20 weeks."

For the purposes of this study, basic data analysis involved comparison of graphed data and trend lines, to aim and goal lines. In this context, trend lines are the line of best fit using ordinary least squares regression, goal lines are the line of performance upon which the decision rule for additional support is predicated (i.e., 30th percentile), and aim lines are lines plotted from a student's/group's baseline performance to the goal line at a future "goal" date. Basic data summarization entails describing: (a) a summary statement of the data for each group relative to both the goal line (level of performance desired) and the aim line (rate of progress needed to reach goal line on the stated timeline), (b) an evaluative statement regarding the presence or absence of a group level problem, and (c) an evaluative statement for each student indicating whether their performance is consistent with the performance of the group or not and summarizing the difference (i.e., "While the group is struggling, Johnny is actually making excellent progress").

Goal setting was discussed in the context of precision problem statements and data analysis. A criterion goal for any given student or group meets the following conditions: (a) match the ‘what’ and ‘who’ of the precision problem statement, (b) have a magnitude of change greater than or equal to the identified ‘to what degree’ portion of the problem statement, (c) include a date by which the identified change will be achieved, and (d) produce a goal that if achieved on the stated timeline will eliminate the disparity between observed and expected performance. Teacher participants were taught to use the ‘what’ and ‘who’ from their precision problem statements in describing their goals for clarity, and to use the goal and aim lines from their data analysis/summarization to generate appropriate magnitudes and timelines for their goal statements. This section ended with a discussion of the confirmation process that the goal meets condition ‘d,’ in which teachers were instructed to check the stated values against the displayed data and goal/aim lines.

During the training on precision problem statements, participants learned to identify with clarity and specificity the critical feature of the problem apparent in the data under discussion. These elements include what problem is present, who is experiencing the problem, to what degree the observed performance differs from what is expected (i.e., “Bobby’s reading fluency is improving, but not fast enough to reach the 30th percentile by the end of the year”), and why it is thought to be occurring (i.e., “We think Bobby is not making fast enough progress because he is not getting enough practice with the intervention curriculum”). For problems with a social behavior component, when and where the problem occurs are also critical components of a precision problem statement, and were covered during this session. Teachers were taught the conditional rule of

including these elements when relevant and omitting them when irrelevant (problems which are solely academic).

Session 3. The third training session began with a review of the previous session's content, with a particular emphasis on design of a comprehensive support plan. New content delivered in Session 3 included selecting intervention solutions matched to precision problem statement, monitoring the impact of implemented solutions, and rendering summative evaluations after solutions have been implemented with fidelity for some period of time. Both of these aspects of problem solving built directly from the integrity with which solutions had been implemented. As such, teachers were taught to first raise the question of whether planned solutions were implemented with integrity, on the pre-specified timeline. If a solution had not been implemented as planned, teams were taught to determine what did not work with the plan and specify a new implementation plan for the pre-selected solution, or a new solution with a corresponding implementation plan. If the implementation plan was simply delayed, but had been in place for some time by the meeting date, review of the outcome data was simply delayed until the next meeting. If the solution had been implemented as planned, teams then analyzed the available data to determine what effect it had on student performance. Teams were taught to describe the effect with a simple statement: the problem has gotten worse, there has been no change, the problem has improved but is not yet resolved, the problem has been resolved, or it is difficult to determine whether the solution has been effective at this time.

Solution identification was closely tied to the discussion of goal-setting and the 'why' of precision problem statements. Training on this portion of the problem solving

process emphasized: selecting solutions from a pre-identified list of malleable factors that the team is charged with modifying as needed, selecting a solution that matches the ‘why’ portion of the stated precision problem statement, selecting a solution for which it is feasible for the group to implement or arrange implementation, and selecting a solution for which the group may reasonably anticipate achievement of the goal by the specified date. Teams were explicitly trained to estimate the magnitude of effect that they may expect for interventions and modifications of varying degrees of intensity. They were taught to use this information to match the intensity of selected solutions to the magnitude and timeline of the goal statement.

Training on implementation planning of solutions was linked to the pre-identified list of malleable factors (which curricula can be used, what grouping decisions can be made, and how many students may be placed into an intervention group). which the group has been charged with modifying as needed. For each solution a team chooses, for any given identified problem, implementation planning in the current context involves identification of several key elements. Key elements of implementation planning include: who will implement the identified solution, when will implementation begin, what data were collected on the fidelity of solution implementation, when will these data be collected, and who will collect these data. Teams were taught to plan for implementation of solutions

Coaching

Coaching is a critical systems-level component in the implementation of evidence-based educational practices (Fixsen, Blase, Naoom, & Wallace, 2009; Fixsen, Naoom, Blase, & Friedman, 2005; Joyce & Showers, 1982; Joyce & Showers, 2002).

Past research has found that without coaching, practices in which teachers are trained demonstrate very low levels of implementation (Joyce & Showers, 2002). Research on implementing the TIPS model for decision making with social behavior data has consistently emphasized the importance of coaching (Horner et al., 2018b; Horner et al., 2009; Newton, Todd, Algozzine, Horner, & Algozzine, 2009; Newton et al., 2011; Newton, Horner, Algozzine, Todd, & Algozzine, 2012b; Todd et al., 2011), consistent with the broader literature within the School-wide Positive Behavioral Supports literature (Horner & Sugai, 2015). Coaching occurred after each training session, but prior to and in the context of regularly scheduled team meetings. Coaching was part of the two meetings immediately following training, but remain available until the team demonstrated proficiency implementing the model. Coaching support was provided by the first author for the first two meetings after each training session and in-between meetings as needed. Check-in coaching sessions occurred with at least one member of the data team five additional times beyond what was planned in advance of the training. The focus of coaching activities was on the content of the most recent training, but for coaching sessions after the second and third trainings, prior training content was coached as needed. Coaching continued until the team demonstrates initial mastery applying the trained content. Initial mastery was determined by the coaches' judgment during the meeting based upon the fidelity with which the team applies the TIPS process. Unprompted success with the component most recently trained prior to the end of the meeting was used as a minimum objective criterion. Coaching activities consisted of prompting preparatory behaviors from members serving in specific roles prior to team meetings to increase the probability of correct application of the TIPS model during

meetings. Additional coaching involving performance feedback and in-session prompting will occur as needed to promote correct application of the TIPS model. Four functions of coaching are critical to implementing complex systems: (a) prompting, (b) fluency-building, (c) performance feedback, and (d) adaptation (Massar & Horner, 2016). The first and third of these are clearly articulated in the role of coaches in the TIPS model, the second and fourth require careful timing of assistance from coaches. To promote fluency, prompting before the initial session following training is necessary to promote accuracy and contextually-situated independent performance. Further, the opportunity to perform as much of each task independently as the team can accurately must be afforded during each coached session. This means application of least to most prompting with contextually appropriate delay of prompting behaviors by coaches to allow sufficient “think time” for teams. Adaptation further requires preparation in advance for anticipated differences between the context of application and traditional contexts of application or study. Additionally, adaptation requires the opportunity for teams to implement TIPS with their best fidelity under contextual constraints, and prompting specific to adaptation as the need for such adaptation becomes clear. This may often occur in the form of problem identification (what constitutes a problem will necessarily depend upon the decision rubric in use by the team) or in the context of comprehensive plan development and contextual fit. Additional considerations regarding adaptation may occur during meetings where the team is reviewing fidelity data on their implementation of the TIPS process.

Session 1. Following the first training session, coaching began one week prior to the DT's next meeting. Pre-session coaching occur via email prompts and face-to-face individual meetings to review preparation and answer questions. Each team member serving a specific role was sent an email prompting them to engage in their meeting preparation, requesting confirmation once preparation tasks are completed, and soliciting any questions on how to exercise the member's role. The facilitator was prompted to review the minutes from the previous meeting, specify the content for the next meeting agenda, allocate meeting time to each portion of the meeting, and disseminate the agenda at least 24 hours prior to the scheduled meeting. The minute keeper was prompted to prepare a meeting minutes form using the agenda shared by the facilitator, pre-filled with available information where possible prior to the meeting (i.e., names of students to be reviewed, names of students/groups for new problem solving, etc.). The data analyst was prompted to review data reports related to new and previous problem solving, and share summary materials with the full team at least 24 hours in advance of the meeting. In-session coaching will focus on least-to-most prompts to increase the fidelity with which each team member engages in their role-related duties.

Session 2. After the second training session, all team members will receive an email summary of the components of precision problem statements and goals at least two days prior to the next scheduled team meeting. Team members were encouraged to ask any questions on these aspects of the problem-solving model prior to the meeting, with additional one-on-one meetings offered as needed. During the meeting, prompting was provided to ensure the facilitator keeps the team focused on the elements of precision problem statements and goals, with some clarification likely provided regarding specific

examples of the different components of precise problem statements and appropriateness of goals (i.e., clarifying how specific a statement of “what” needs to be, how much progress can reasonably be expected). Additional prompting, clarification, and re-direction regarding basic data analysis will occur as needed to ensure that teams are accurately analyzing their students’ data.

Session 3. Following the third training session, all team members received an email summary of solution identification (emphasizing the link to goal statements and the “why” portion of the precisely identified problem). An email prompt was sent to the facilitator prompting follow-up on any implementation monitoring tasks that was assigned to members of the team. In-session coaching will emphasize supporting the facilitator in appropriately constraining conversation regarding solution identification and implementation to malleable factors which are linked to specific aspects of the precision problem statement or goal. Additionally, prompting to engage in summative evaluation decisions for past problems will occur as needed.

Design and Analysis

This study utilized a concurrent, multiple baseline across skills design to answer the first research question. Trainings occurred three times during the school year, once each for the three different training content areas. The core team (DT) training events were scheduled within each training phase based upon the conditions necessary for a multiple baseline across skills design. The trainings were further scheduled with the needs and availability of the school team members in mind. For this analysis, the unit of analysis was the core data team (DT) and participating GLTT members. The facilitator

was the same core team member for each iteration of the data team. Additional team members filled the remaining roles.

Analysis of the DORA data within the multiple baseline design was done via visual analysis. Visual analysis of multiple baseline data involves two levels of comparison (Kratochwill et al., 2010). The first is a standard evaluation for basic effect within each unit (Foundations, Decision Making, Solution Implementation). The essential aspects of graphed data appraised in this analysis are changes in: level, variability, and trend. Further, the immediacy of effect is a critical factor in analyzing such data, with more immediate effects being considered more compelling demonstrations of experimental control (particularly in the absence of a compelling theoretical explanation for why one should expect to see delayed effects). Finally, multiple baseline data are analyzed for consistency across similar phases. For a multiple baseline across skills (or across behaviors) design, the absolute level of performance prior to intervention is not hypothesized to be similar in all cases. As such, it is the post-training performance that is of greatest interest for this comparison for this particular single case design. It is important for confident interpretation that responding in each phase achieves a reasonable approximation of the “steady-state,” wherein the performance on the dependent variable is at a reasonably consistent level within each phase prior to moving onto the next phase. For highly variable behavior, or for applications of the multiple baseline design which preclude fully response-guided intervention timing, a clear and understandable pattern of responding may be acceptable for visual analysis to proceed with confidence.

The second research question was addressed using descriptive quantitative analyses. EasyCBM data were collected for all students in the school three times per year

for benchmarking (fall, winter, and spring) as part of the school's standard operating procedures. The proportion of students at each level of risk (low, some, or high) was calculated for each grade level, during each benchmarking period. A greater proportion of students at low risk indicates a better functioning system of reading supports. A trend across the academic year of more students achieving lower risk status indicates more effective decision making and delivery of supplementary supports (Tiers II & III).

The third research question was addressed using survey responses from the members of the DT. The five core DT members each completed two surveys: the Adapted Self-Assessment of Contextual Fit, and the Primary Intervention Rating Scale. Both surveys present a set of items (16 & 17 respectively) about the perceived effectiveness and acceptability of a given intervention. Respondents are instructed to rate their level of agreement with the statement provided by each item on a Likert scale from 1-6. The PIRS also includes four open-ended responses. Surveys were analyzed for mean rating by section, with representative quotations provided from the open-ended responses.

Effect Size

Non-overlap of all pairs (NAP) was used to calculate the size of the effect for the multiple baseline design examining the effect of team training on implementation of TIPS procedures for academics (Parker & Vannest, 2009). NAP is calculated using the R package SingleCaseES (Pustejovsky, 2017). NAP estimates the size of an effect for single case data (including MBL) using the probability that a randomly selected treatment phase data point will exceed a randomly selected baseline phase data point. Non-overlap of all pairs is a measure of the degree to which values in each phase are unique to that phase. It is a measure of the amount of non-overlapping data relative to all possible

comparisons between phases. It is closely related to the common language effect size and identical to the probability of superiority (Parker, Vannest, & Davis, 2011). NAP is interpreted as the probability that a randomly selected data point from the post-training phase will be superior to (in this case higher than) a randomly selected point from baseline. NAP scores naturally range from .5 to 1.0 (50% to 100%), where .5 is chance-level and higher scores indicate a larger effect. NAP can be adjusted to a 0 to 1 scale, with zero being equal to chance level probability and higher values indicated a stronger effect of intervention on a scale more similar to other measures of effect size.

CHAPTER IV

RESULTS

The purpose of the present study was to evaluate the effects of professional development in Academic Team-initiated Problem Solving (AcTIPS) on the implementation of evidence-based decision-making practices by an elementary literacy data team. A concurrent multiple baseline across skills design was used to assess the team's implementation of core components of the AcTIPS model before and after training. Training occurred in three separate sessions, beginning with meeting foundations, proceeding with problem-solving, and concluding with solution implementation and adaptation. Results are summarized for (a) direct observation of team meetings within a multiple baseline, (b) student literacy outcomes, and (c) staff perceptions of the social validity of AcTIPS training and procedures.

Direct Observation Data

Data from direct observation of team meetings is provided in Figure 1. During the baseline phase the team demonstrated low use of core TIPS procedures. The seven baseline data points for Meeting Foundations averaged 41.4% with a range of 40% to 50%, and modest trend. The twelve baseline data points for Problem Solving were more variable with a mean of 63% and a range from 44% to 89%, with no meaningful trend. The fifteen baseline data points for Solution Implementation and Adaptation indicate this content was the least well performed by the team, with a mean of 27%, a range of 0% to 50% with no clear trend. Prior to AcTIPS training, the team consistently omitted the “why” portion of their problem statements and omitted another (what or who) sometimes

as well. During the baseline phase the team included all three elements for a single precision problem statement for 19% of identified problems (5 of 26). The team included “what” 58% of the time prior to training, “who” 88%, and “why” 19%. When the team was able to get “why” into their conceptualization of a given problem, they always had the other two components as well (5 of 26 identified problems). Additionally, goal statements were missing the magnitude, timeline, or both elements consistently prior to training as well. Magnitude was included for 23% of observations and timeline for 19%. Both elements of a criterion goal statement were present for 15% of problems prior to training. All aspects of problem solving (including precision statements, goals, and usage of quantitative data) were included for 4% of problems during baseline. Solution implementation and adaptation elements were observed with similar infrequency prior to training. For newly identified problems, an individual was identified to execute changes the team decided upon 92% of the time during baseline (12 of 13 newly identified problems). A timeline for implementation of a selected intervention was included 31% of the time (4 of 13 new problems), and the team never planned any type of fidelity measure or reporting during baseline. For pre-defined problems, implementation integrity for interventions previously selected was never reported, the status of the problem (current student performance) was reported 67% of the time (8 of 12 old problems), comparison between this performance and the stated goal was made 25% of the time (3 of 12 old problems), and a summative evaluation decision was rendered 50% of the time (6 of 12 old problems). The team never incorporated all elements for this phase of problem solving during baseline for either new or old problems.

Following implementation of AcTIPS training the team improved use of effective team meeting procedures. For Meeting Foundations, the team demonstrated an immediate and sustained improvement following training. The average score for Meeting Foundations across the 14 team meetings following training was 86.4% with a range of 70% to 100% and no trend. The effect size as assessed by NAP was 100%.

A similar pattern was observed with Problem Solving. There was an immediate and sustained improvement in the team's use of problem solving procedures following AcTIPS training. The mean performance on Problem Solving after training was 90% with a range from 78% to 100%, a reduction in variability, and a slight increasing trend across the eight team meetings following training. NAP for implementation of Problem-Solving skills is 94.3% with a standard error of .047. This indicates that the probability of problem-solving performance occurring during a randomly selected meeting after training exceeding the problem-solving performance during a randomly selected meeting occurring prior to training is 94.3%. In terms of problem precision, statements created following AcTIPS training included all critical elements for 90% of problems identified by the team. Statements included "what" 100%, "who" 100%, and "why" 90% of the time following training. Goal statements were also substantively improved, with 50% including both elements, 60% including magnitude, and 50% including a timeline for achieving the stated goal. The goal statement portion was significantly more challenging for the team, and it is worth noting that the distribution of goal statements across meetings was not evenly distributed after training. Rather, the last five team meetings included both aspects of goal statements 100% of the time. Additionally, all aspects of

problem solving with precision were present in 40% of problems identified after training, again with these problems occurring during four of the last five team meetings.

The results for Solution Implementation and Adaptation skills following AcTIPS training also indicate an immediate and substantive effect. Results for Solution Implementation and Adaptation scores from the five team meetings following training averaged 90% with a range of 75% to 100%, and a NAP of 100%. There were no newly identified problems after training this portion of the AcTIPS training due to the time of year at which it occurred. For previously identified problems, implementation integrity for interventions previously selected was reported 60% of the time, the status of the problem (current student performance) was reported 100% of the time, comparison between this performance and the stated goal was made 100% of the time, and a summative evaluation decision was rendered 100% of the time. This yielded complete implementation of this portion of the TIPS model 60% of the time following AcTIPS training.

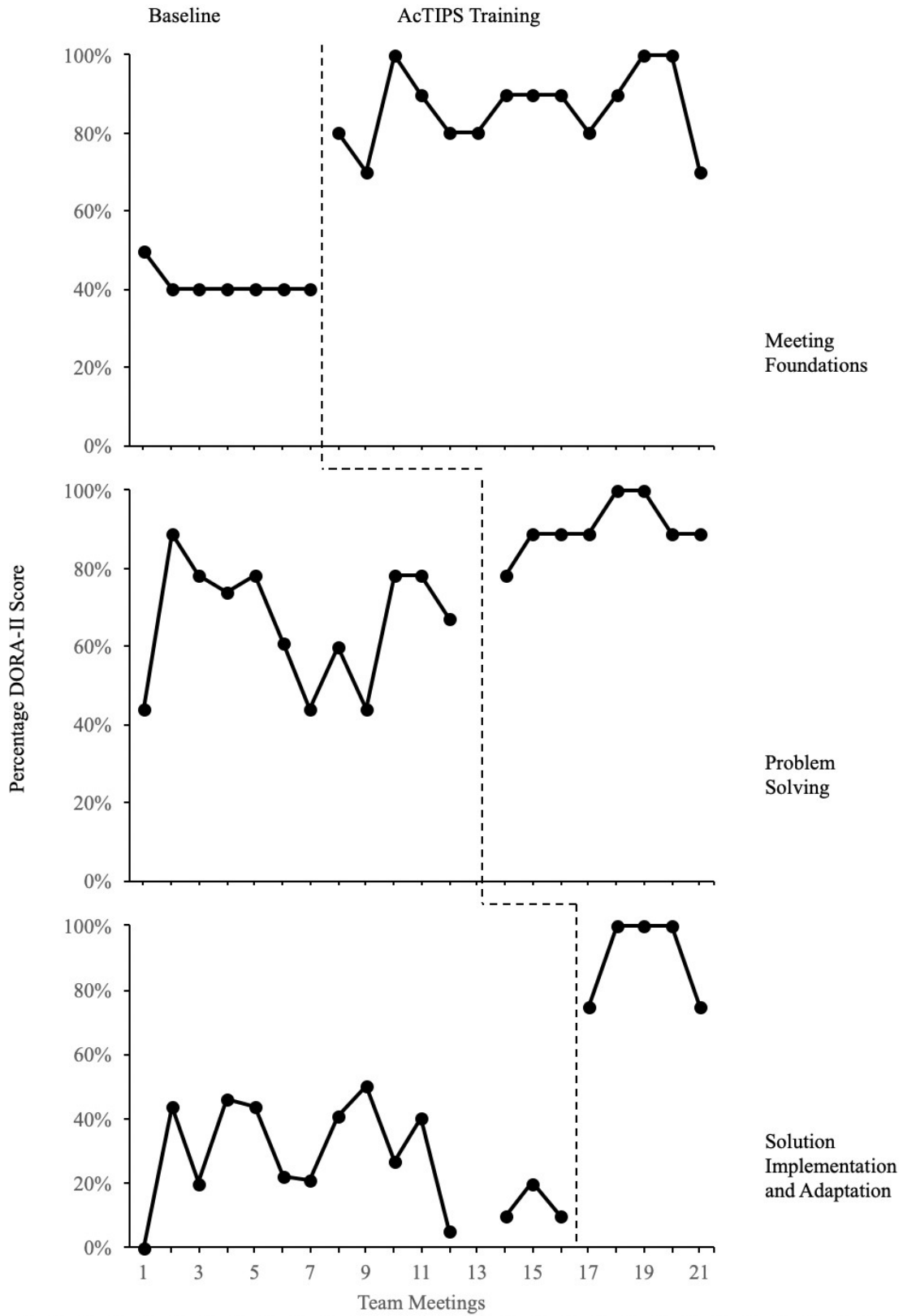


Figure 1. DORA-II Scores Before and After AcTIPS Training

Student Literacy Outcomes

Student literacy outcomes for the present study were evaluated using a descriptive comparison of the students at elevated risk for reading failure during the benchmarking period immediately prior to training and after all training had been delivered during the academic year of the study. For comparison, the proportion of students at elevated risk during the corresponding benchmarking periods during the prior two years is also presented. Summary of these data are presented in Table 1 below.

For the 2018-19 school year, students at elevated risk for reading failure ranged from a low of 13% of Kindergarteners after training to a high of 42% of second graders prior to training. Prior to AcTIPS training the proportion of students at elevated risk was 34% in Kindergarten, 30% for first grade, 42% for second grade, 31% for the combined 3rd/4th grade group, and 23% for grade 5. The next benchmarking period following the completion of all training, the proportion of students at elevated risk for reading failure was 13% in Kindergarten, 35% in first grade, 40% in second grade, 37% in the combined 3rd/4th grade group, and 27% in fifth grade.

For the 2017-18 and 2016-17 academic years, student risk for comparable benchmarking periods ranged from a low of 15% for second graders in the spring of 2017, to a high of 46% for second graders in the winter of 2018. For the 2016-17 school year, data were not collected for Kindergarteners or first graders. Proportion of students at elevated risk for other grades during winter of 2017 was 30% for second graders, 26% for third graders, 22% for fourth graders, and 27% for fifth graders. During spring of 2017, these values changed to 15% for second graders, 32% for third graders, 30% for fourth graders, and 24% for fifth graders.

For the 2017-18 school year, no data were collected for Kindergarteners. Proportion of students at elevated risk during winter 2018 for other grades was 32% for first graders, 46% for second graders, 36% for third graders, 29% for fourth graders, and 37% for fifth graders. During spring of 2018, the proportion of students at risk for grades with data collected was 25% of first graders, 44% of second graders, 29% of third graders, 28% of fourth graders, and 28% of fifth graders. The AcTIPS training package was delivered from February through May of the school year, and there was simply not enough time to see results from changes made so late in the year.

Table 1. Proportion of Students at Elevated Risk of Reading Failure

	K	1	2	3	4	5
Year	W / S	W / S	W / S	W / S	W / S	W / S
16-17	NC/NC	NC/NC	30%/15%	26%/32%	22%/30%	27%/24%
17-18	NC/NC	32%/25%	46%/44%	36%/29%	29%/28%	37%/28%
18-19	34%/13%	30%/35%	42%/40%	31% /	37%	23%/27%

Perceptions of Social Validity

Two surveys were administered to members of the core DT: the Primary Intervention Rating Scale (PIRS), and the Adapted Self-Assessment of Contextual Fit (ASACF). Ratings for items on the PIRS ranged from 4 to 6, with the exception of the item pertaining to the similarity of AcTIPS to other models of decision making with which staff had experience. This item was rated a 3 by one member of the core DT

indicating that this individual perceived TIPS more different than similar to other models of decision making. Mean ratings on the PIRS was 5.02, indicating that AcTIPS was highly acceptable and valued by members of the core DT as measured by PIRS ratings. Similarly, item ratings for the ASACF ranged from 4 to 6. Mean rating for the ASACF was 5.15, providing convergent evidence that the AcTIPS model of decision making and training were valued and acceptable to core DT members. Correlation between these two measures of acceptability was .65, indicating that the two surveys are moderately positively related to one another even in such a small sample (Table 2).

Table 2. Perceptions of the Social Validity of AcTIPS Training and TIPS Implementation.

Respondent	PIRS	ASACF
1	85	79
2	85	80
3	83	81
4	87	83
5	87	89

CHAPTER V

DISCUSSION

This study focused on the need to provide school teams with protocols for using behavioral and academic data in regular decision-making. While data-based decision-making has become a hallmark of effective education, more attention has been paid to the development and collection of measures (both academic and behavioral) than to the process teams use to make decisions from the resulting data. Recent results from the Team Initiated Problem Solving approach with behavioral data have been encouraging, and the current research sought to determine if that framework could be trained, used, and effective with elementary school teams focused on literacy outcomes in elementary schools.

Summary of Findings

The overall results of the present study are positive, indicating that AcTIPS training was both acceptable to members of an elementary literacy data team, and functionally related to increased use of evidence-based decision-making practices. The data on student outcomes indicate that implementing the full AcTIPS model by the end of May did not impact student performance on spring benchmarking. This is to be expected, but clarifies the need to complete the full AcTIPS training earlier in the year in future studies so that sufficient time implementing the model can pass for the effect of implementing AcTIPS on student reading outcomes can be clearly evaluated. The clear change in level immediately following training associated with each of the three skills targeted indicates a functional relation between AcTIPS training and implementation of

data-based decision-making practices. Further, the present study utilized a multiple baseline design with at least five data points per phase thus meeting the requirements for the full What Works Clearinghouse standards for single case designs (Kratochwill et al., 2010). Additionally, the study demonstrated changes in performance as a function of AcTIPS training at three distinct points in time, meeting the criterion for documentation of an experimental effect. Finally, the results of the present study provide an initial demonstration of the utility of the multiple baseline across skills design, the analog of a multiple baseline across behaviors design applicable to research utilizing teams or groups as the unit of analysis.

Team-based Decision Making

Implementing team meeting foundations is critical to conducting effective data-based decision-making to improve student outcomes. Prior to training, the DT was implementing approximately half of the research-identified core aspects of effective team meetings (40-50%). Once the DT was trained in the roles associated with the TIPS framework, a small but noticeable increase in problem-solving performance occurred. While this increase was not enough to achieve the desired performance criterion, it does reinforce the basic logic of the TIPS model which is predicated on the efforts of team members filling roles completing specific aspects of the problem-solving process. Although the training on roles omitted any material which would overtly be associated with problem-solving, the establishment of roles and the basic understanding of what was expected of each member was enough to help the team improve its performance in this critical domain. It is possible, that team members had some idea of how to complete the various tasks associated with data-based decision making prior to the first training, but

that a sort of “diffusion of responsibility” resulting from a lack of clearly defined roles rendered them less likely to complete these aspects. The AcTIPS training then may have provided the necessary clarification of expectations for team members to take ownership of the process and engage their best performance. The improvement in implementing team meeting foundations is at once clear, immediate, and sustaining. This indicates that the team did not have significant challenges implementing these aspects of the model once they were explained in a systematic training session and subsequent coaching was provided.

Building on this result are the findings that “why” elements were most challenging for the team and when this aspect of a precision problem statement was included, all others were very likely to be included as well. Additionally, a similar observation is clear related to the timeline aspect of the goal statement, planning for treatment integrity during intervention selection, and reporting on treatment integrity at follow-up. Each of these elements was clearly more challenging for the team and when these aspects of each domain were present, the others were much more likely to be present as well. These items may be useful for inclusion in a short job aide, or for building a briefer self-assessment tool. Additional research is needed to determine the generalizability of these patterns and their utility for such purposes.

The improvement in implementing the problem-solving portion of the TIPS decision-making model is similarly compelling in its immediacy and sustained improvement. The variability observed during the baseline condition of this skill (prior the AcTIPS training on problem solving) was significant. The real-world impact of the changes observed on this skill are meaningful. The combination of increased level and a

large reduction in variability means that the problems this team solved after training were much more likely to include the critical components necessary to match student needs to appropriate interventions. Highly variable problem-solving effectiveness is itself problematic. For school teams to make decisions that reliably benefit students they must engage a clear and effective data-based decision-making process consistently. One may consider what would occur if such a team made excellent decisions every other month and poor decisions on the off months. Students would benefit for a short period of time from well-conceived interventions and instructional programs once implemented following the effective meetings, and some of that progress would surely be undone during the less effective meetings. In practice, problem-solving needs to be effective consistently because that is the only efficient option. There are simply too many children who need intervention decisions in a school of 485 enrollment with 30-40% at elevated risk depending on the grade level for some meeting minutes to be spent ineffectively. At fall benchmarking the year of the study, 37% (116) of the 444 students who were administered benchmarking assessments were at elevated risk. Staff at the participating school are now much more likely to identify student reading problems with sufficient precision as to be able to match student needs to appropriate interventions.

The third skill set trained was Solution Implementation and Adaptation.

Following this training and related coaching there was an immediate increase in the level of team performance in this area of decision making. The change in performance is both clear and sustained over the duration of the study. Given the low level of attention given to this portion of the decision-making process by the DT during baseline, it is not surprising that training had a large and meaningful effect. The impact of changes in this

portion of the decision-making process should not be understated however. When teams do not make evaluative judgments about what interventions or intervention components have or have not worked for a student, it is not possible to ensure students get their needs met over time. Implementation of this component of data-based decision-making is critical to sustained implementation of the broader model as well. Failure to implement this part of the model yields conditions wherein students who are not matched with the right intervention on the first pass will not be matched with effective supports during their time at the school in question. The core purpose of implementing data-based decision-making is to improve student outcomes and when teams fail to do so (or do not come into contact with confirmation of their success) they may likely be inclined to persist less with the components that are working simply due to the lack of effective systemic feedback on their performance. Thus, the large improvement in this domain of decision-making is important as it sets the necessary conditions in place for more consistent improvements in student behavior and sustained implementation of an evidence-based decision-making model. Across all three TIPS skills trained in the present study, significant improvements were observed following training. The impact of the observed changes in implementing the TIPS model of data-based decision-making mean that students at the participating school are now much more likely to benefit from effective problem-solving regarding their reading performance.

Student Outcomes

While it is always critical to gather data on student outcomes, the present study is purely descriptive in terms of the student outcome data included. The data indicate that no meaningful change occurred in student outcomes relative to typical intra-year changes

in performance observed in prior years during the limited period of the academic year following completion of the full AcTIPS training. It is not possible to know what effect AcTIPS professional development and implementation of the TIPS model for literacy decisions may have on students' reading outcomes from the present study. The full TIPS model was not implemented until late in the year and the mechanism by which data-based decision-making impacts student behavior requires implementation of the model for a sufficient amount of time that better intervention decisions are made, better interventions are implemented, and these superior interventions have time to exert a stronger positive effect on student performance than less well-matched interventions. A detailed analysis using precise identification of individual students may achieve this in as little as three months for reading interventions. A group level analysis of student reading performance like that used in the present study would likely require nearly an entire school year of full implementation (perhaps longer). Thus, the few weeks at the end of the year were simply not enough time to associate changes in student behavior with implementation of the AcTIPS model.

Social Validity

The core data team appreciated the training, felt it was useful, and found it effective in helping them use their data more effectively. The overall mean for the two surveys both exceeded 5 on a scale from 1-6, indicating that the team was strongly supportive of the training and implementation of the model. In particular, several team members made efforts to highlight their perception of the positive impact AcTIPS was having on their students' reading instruction. In response to the question of whether AcTIPS had resulted in improvements to their students reading performance and reduced

reading problems, staff responded with comments like, “I think that it helped us take a closer look at our students’ skills and make adjustment to better meet their needs,” and “Yes, it has been instrumental to our implementation of data teams and RTI.” Members of the DT also consistently noted the improvement they perceived in their usage of data and expressed how much they appreciated the training session focused on data analysis and precision problem statements. Average ratings for individual items across both surveys ranged from 4.4 to 5.6, all within a range that could be considered solidly endorsing the acceptability of the model and procedures used for training.

Research Design

The present study utilized a multiple baseline across skills design, with a team of educators as the unit of analysis. This design is analogous to the well-established multiple baseline across behaviors design used with individuals (Bailey & Burch, 2002; Cooper, Heron, & Heward, 2007). The results indicate that the logic of the multiple baseline across behaviors design does indeed transfer well to the study of teams or groups of individuals. This is clear in the vertical analysis comparing the relative stability of and level of different skills across phases. Changes in team performance of skills occurred as a function of training and the analysis is straightforward in keeping with analysis of data from the analog design as used with individuals. One point of interest here is that a modest increase occurred in the team’s execution of the problem-solving skill following their implementation of roles (and suffered when their implementation of roles was less complete). This is analogous to the situation of concern with this design when used with individuals wherein an individual may generalize learning related to one behavior to other behaviors. When properly accounted for and documented, this is actually a strength

of this design as it allows for the partial evaluation of such generalization effects. In the present study, it indicates that implementation of TIPS roles may garner some level of problem-solving benefit all on their own. Finally, while this created some variability that was not due to the training targeting this skill, the design itself was robust to this challenge. This demonstrates that this design is suitable for application to contexts wherein skills are generally and predominantly conceptually independent of one-another, without an expectation that they be absolutely independent of one-another.

Limitations

The present study is limited by the inclusion of a single data team in the training and observations. Generalization of these findings to other teams requires significant caution. Further, as this research was conducted in an elementary school in Western Oregon, the generalization of these findings to other grade levels of teams in other regions requires replication with a broader sample. Additionally, the attenuated time frame over which this study took place, six months, calls for caution when interpreting effects beyond those directly associated with training the data team. The timing of the training and size of the sample, as well as fundamental features of the research design utilized in the present study rendered all analysis of student outcomes descriptive. As such, it is worth emphasizing that the focus of the study was on the impact of AcTIPS training on the behavior of the core team, and it is not reasonable to draw conclusions about the effect of AcTIPS professional development on student outcomes at this time.

Implications and Future Directions

While the results of the present investigation are promising regarding the value of the AcTIPS professional development for supporting data teams in implementing evidence-based data-based decision-making practices, additional work needs to be done in this area. First, while the internal validity of the present study is strong, the use of one team in a single case design indicates a need for replication of the training procedures with additional teams to support stronger external validity. Second, the potential of the TIPS model to be applied productively to student challenges relating to both social and academic behavior concurrently is an area of obvious potential value. The overlap of systems involved in RTI and PBIS is undeniable and the need to examine the performance across both social and academic behavior for some children is clear. One need only consider a child with escape-maintained problem behavior as a function of low reading skill to see the value of integrated decision-making frameworks. Third, the specific factors that may drive sustained implementation of AcTIPS following training are as yet still unclear. This is an area of clear need for further study. Further, as noted previously, the emergence of clear patterns of more challenging sub-skills has important implications for both training in data-based decision-making and assessment thereof. As the field of curriculum-based measurement has made clear, highly efficient measurement is possible when one focuses on the meaningful and challenging composite skill of interest. Part of the reason for this efficiency is that the nature of a composite task is to recruit the performance of the relevant component repertoires and thus measurement of the composite serves as a fair proxy for proficiency with the underlying skills as well. The analog present here is that a clear “why” element may serve as such an index of

precision statement complexity. Similarly, inclusion of a timeline, plan for treatment integrity, and reporting of treatment integrity, may be indicators of proficiency for the other critical stages of the problem-solving process. Finally, the impact of AcTIPS implementation for a full school year on student outcomes warrants further examination as the present study did not allow for sufficient time implementing the full model for effects to be discerned.

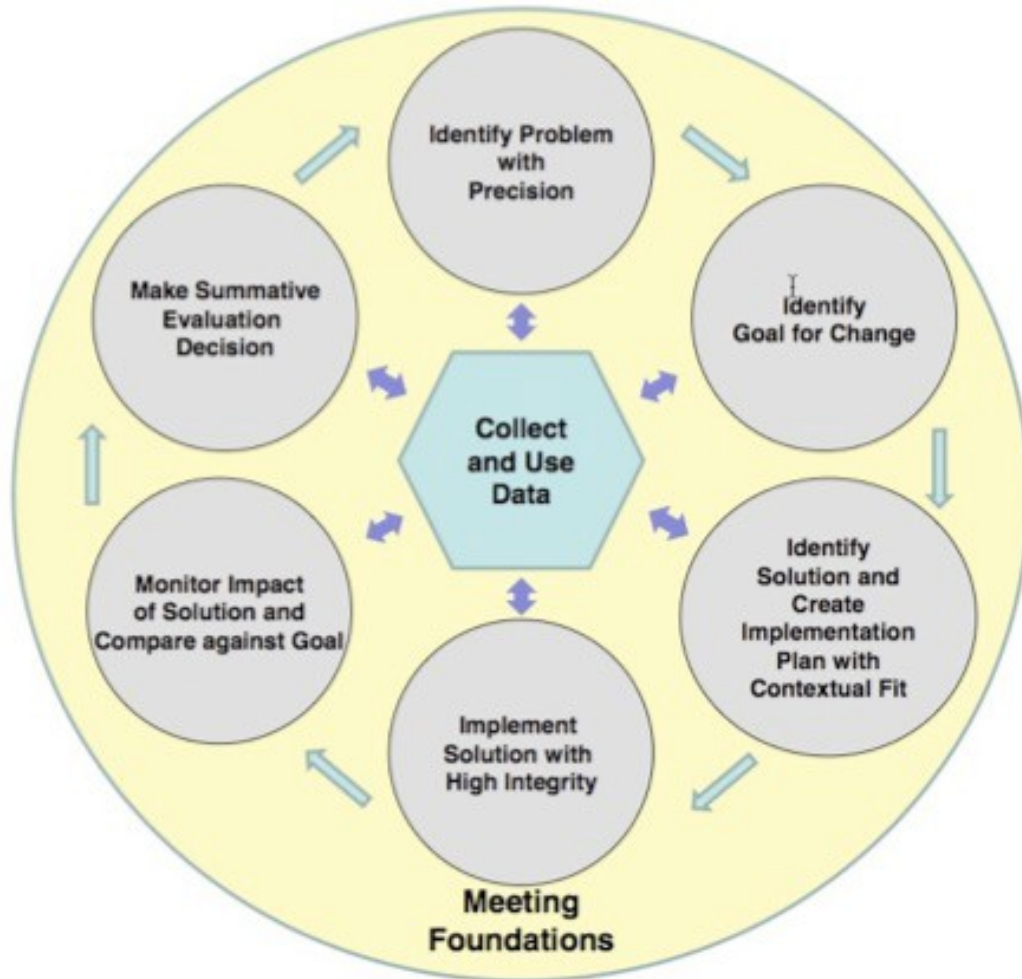
APPENDIX A

AcTIPS Professional Development



APPENDIX B

AcTIPS Training Manual



**Academic Team-Initiated Problem Solving
Training Manual**

APPENDIX C

Decision Observation, Recording and Analysis II

Section 1. Demographic Information

School ID No.: _____ # PBIS Team Members: _____ Observer Name: _____ Primary Observer Reliability Observer

Date: _____ Scheduled Start Time: _____ Scheduled End Time: _____

If these are research data complete the following:
 Group: _____ State: _____ Condition: _____ Data Wave No.: _____

Section 2. Foundations of Effective Team Problem Solving

START OF MEETING	DURING MEETING (ROLES)	END OF MEETING
01. <input type="checkbox"/> Meeting started within 10 minutes of scheduled start time	05. <input type="checkbox"/> Facilitator	08. <input type="checkbox"/> Next meeting scheduled
02. <input type="checkbox"/> At least 75% of team members present at the start of the meeting	06. <input type="checkbox"/> Minute Taker	09. <input type="checkbox"/> Meeting ended within 10 minutes of scheduled end time (includes a revised end time that team members agreed to)
03. <input type="checkbox"/> Previous meeting minutes available	07. <input type="checkbox"/> Data Analyst	10. <input type="checkbox"/> At least 75% of team members present at the end of the meeting
04. <input type="checkbox"/> Agenda available		

Notes:

2013-02-20 - DORA II Page No.

Section 3. Team Problem-Solving Processes

Operational definition of a "problem" - At least one team member or meeting participant identifies a student social or academic behavior to change, AND the team selects/selected a solution to bring about the desired change.

PRECISE PROBLEM & GOAL FOR CHANGE IDENTIFIED	SOLUTION IDENTIFIED AND IMPLEMENTATION PLAN CREATED & SOLUTION IMPLEMENTED WITH INTEGRITY	IMPACT OF SOLUTION MONITORED AND COMPARED AGAINST GOAL & SUMMATIVE EVALUATION DECISION MADE																																																																																																																																																																																																																									
<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 20%;">Problem No.</td> <td style="width: 20%;"></td> <td style="width: 20%;">PR</td> <td style="width: 20%;"></td> </tr> <tr> <td colspan="4">Description of identified problem</td> </tr> <tr> <td colspan="4" style="height: 50px;"></td> </tr> <tr> <td colspan="4">1.1 Problem Precision</td> </tr> <tr> <td>What</td> <td>Who</td> <td>Where</td> <td>When</td> </tr> <tr> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td colspan="4">Problem Category</td> </tr> <tr> <td colspan="2">Social Behavior <input type="checkbox"/></td> <td colspan="2">Academic Behavior <input type="checkbox"/></td> </tr> <tr> <td colspan="4">Problem Features</td> </tr> <tr> <td>New</td> <td>Old</td> <td>Individual</td> <td>Group</td> </tr> <tr> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td colspan="4">1.2 Quantitative Data Use</td> </tr> <tr> <td colspan="2">Social Behavior <input type="checkbox"/></td> <td colspan="2">Academic Behavior <input type="checkbox"/></td> </tr> <tr> <td colspan="4">Description of data presented</td> </tr> <tr> <td colspan="4" style="height: 50px;"></td> </tr> <tr> <td colspan="4">2. Identified Goal</td> </tr> <tr> <td>What Change</td> <td>By</td> <td>When</td> <td></td> </tr> <tr> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td></td> </tr> <tr> <td colspan="4">Description of change to be achieved</td> </tr> <tr> <td colspan="4" style="height: 50px;"></td> </tr> <tr> <td colspan="4"><input type="checkbox"/> Postponed/out of time</td> </tr> <tr> <td colspan="4"><input type="checkbox"/> Old problem not discussed</td> </tr> </table>	Problem No.		PR		Description of identified problem								1.1 Problem Precision				What	Who	Where	When	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Problem Category				Social Behavior <input type="checkbox"/>		Academic Behavior <input type="checkbox"/>		Problem Features				New	Old	Individual	Group	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	1.2 Quantitative Data Use				Social Behavior <input type="checkbox"/>		Academic Behavior <input type="checkbox"/>		Description of data presented								2. Identified Goal				What Change	By	When		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		Description of change to be achieved								<input type="checkbox"/> Postponed/out of time				<input type="checkbox"/> Old problem not discussed				<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td colspan="6">Description of selected solution</td> </tr> <tr> <td colspan="6">Write description of Solution below, including its individual components or "solution actions."</td> </tr> <tr> <td colspan="6" style="height: 100px;"></td> </tr> <tr> <td colspan="6">3. Solution Implementation Plan</td> </tr> <tr> <td>NA</td> <td>Old</td> <td>Imp.</td> <td>Integrity</td> <td>When</td> <td>Integrity</td> </tr> <tr> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td colspan="6">Description of implementation integrity plan</td> </tr> <tr> <td colspan="6" style="height: 50px;"></td> </tr> <tr> <td colspan="6">4. Solution Implementation-Integrity</td> </tr> <tr> <td>NA</td> <td>Not Started</td> <td>Part. Imp.</td> <td>Imp. w/ Integrity</td> <td>Stopped.</td> <td>Not reported/DK</td> </tr> <tr> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td colspan="6"><input type="checkbox"/> Postponed/out of time</td> </tr> </table>	Description of selected solution						Write description of Solution below, including its individual components or "solution actions."												3. Solution Implementation Plan						NA	Old	Imp.	Integrity	When	Integrity	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Description of implementation integrity plan												4. Solution Implementation-Integrity						NA	Not Started	Part. Imp.	Imp. w/ Integrity	Stopped.	Not reported/DK	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/> Postponed/out of time						<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td colspan="3">5.1 Status of Problem Reported</td> </tr> <tr> <td>NA New Problem</td> <td>Yes</td> <td>No</td> </tr> <tr> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td colspan="3">Description of status of problem (i.e., summary of findings from qualitative and/or quantitative data) or NA if New Problem</td> </tr> <tr> <td colspan="3" style="height: 50px;"></td> </tr> <tr> <td colspan="3">Data source for report on status of problem (NA if New Problem)</td> </tr> <tr> <td colspan="3" style="height: 50px;"></td> </tr> <tr> <td colspan="3">5.2 Status of Problem Compared against Goal</td> </tr> <tr> <td colspan="3">(Note: Check "No" if team has not identified a goal)</td> </tr> <tr> <td>NA New Problem</td> <td>Yes</td> <td>No</td> </tr> <tr> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td colspan="3">6. Summative Evaluation Decision</td> </tr> <tr> <td colspan="3">(Note: Check "No" if team did not report status of problem)</td> </tr> <tr> <td colspan="3">(Note: Examples of summative evaluation decisions include (a) retaining, revising, or terminating (a) the solution, (b) the goal, (c) the precisely-defined problem, or (d) some combination of the preceding)</td> </tr> <tr> <td>NA New Problem</td> <td>Yes</td> <td>No</td> </tr> <tr> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> <td><input type="checkbox"/></td> </tr> <tr> <td colspan="3">Description of decision or NA if New Problem</td> </tr> <tr> <td colspan="3" style="height: 50px;"></td> </tr> <tr> <td colspan="3"><input type="checkbox"/> Postponed/out of time</td> </tr> </table>	5.1 Status of Problem Reported			NA New Problem	Yes	No	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Description of status of problem (i.e., summary of findings from qualitative and/or quantitative data) or NA if New Problem						Data source for report on status of problem (NA if New Problem)						5.2 Status of Problem Compared against Goal			(Note: Check "No" if team has not identified a goal)			NA New Problem	Yes	No	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	6. Summative Evaluation Decision			(Note: Check "No" if team did not report status of problem)			(Note: Examples of summative evaluation decisions include (a) retaining, revising, or terminating (a) the solution, (b) the goal, (c) the precisely-defined problem, or (d) some combination of the preceding)			NA New Problem	Yes	No	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Description of decision or NA if New Problem						<input type="checkbox"/> Postponed/out of time		
Problem No.		PR																																																																																																																																																																																																																									
Description of identified problem																																																																																																																																																																																																																											
1.1 Problem Precision																																																																																																																																																																																																																											
What	Who	Where	When																																																																																																																																																																																																																								
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																																																																																																																																																																																																								
Problem Category																																																																																																																																																																																																																											
Social Behavior <input type="checkbox"/>		Academic Behavior <input type="checkbox"/>																																																																																																																																																																																																																									
Problem Features																																																																																																																																																																																																																											
New	Old	Individual	Group																																																																																																																																																																																																																								
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																																																																																																																																																																																																								
1.2 Quantitative Data Use																																																																																																																																																																																																																											
Social Behavior <input type="checkbox"/>		Academic Behavior <input type="checkbox"/>																																																																																																																																																																																																																									
Description of data presented																																																																																																																																																																																																																											
2. Identified Goal																																																																																																																																																																																																																											
What Change	By	When																																																																																																																																																																																																																									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																																																																																																																																																																																																									
Description of change to be achieved																																																																																																																																																																																																																											
<input type="checkbox"/> Postponed/out of time																																																																																																																																																																																																																											
<input type="checkbox"/> Old problem not discussed																																																																																																																																																																																																																											
Description of selected solution																																																																																																																																																																																																																											
Write description of Solution below, including its individual components or "solution actions."																																																																																																																																																																																																																											
3. Solution Implementation Plan																																																																																																																																																																																																																											
NA	Old	Imp.	Integrity	When	Integrity																																																																																																																																																																																																																						
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																																																																																																																																																																																																						
Description of implementation integrity plan																																																																																																																																																																																																																											
4. Solution Implementation-Integrity																																																																																																																																																																																																																											
NA	Not Started	Part. Imp.	Imp. w/ Integrity	Stopped.	Not reported/DK																																																																																																																																																																																																																						
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																																																																																																																																																																																																						
<input type="checkbox"/> Postponed/out of time																																																																																																																																																																																																																											
5.1 Status of Problem Reported																																																																																																																																																																																																																											
NA New Problem	Yes	No																																																																																																																																																																																																																									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																																																																																																																																																																																																									
Description of status of problem (i.e., summary of findings from qualitative and/or quantitative data) or NA if New Problem																																																																																																																																																																																																																											
Data source for report on status of problem (NA if New Problem)																																																																																																																																																																																																																											
5.2 Status of Problem Compared against Goal																																																																																																																																																																																																																											
(Note: Check "No" if team has not identified a goal)																																																																																																																																																																																																																											
NA New Problem	Yes	No																																																																																																																																																																																																																									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																																																																																																																																																																																																									
6. Summative Evaluation Decision																																																																																																																																																																																																																											
(Note: Check "No" if team did not report status of problem)																																																																																																																																																																																																																											
(Note: Examples of summative evaluation decisions include (a) retaining, revising, or terminating (a) the solution, (b) the goal, (c) the precisely-defined problem, or (d) some combination of the preceding)																																																																																																																																																																																																																											
NA New Problem	Yes	No																																																																																																																																																																																																																									
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																																																																																																																																																																																																																									
Description of decision or NA if New Problem																																																																																																																																																																																																																											
<input type="checkbox"/> Postponed/out of time																																																																																																																																																																																																																											

APPENDIX D

Primary Intervention Rating Scale

Name: _____ Date: _____

School: _____ County: _____

Primary Intervention Rating Scale: Post-Implementation Teacher Version – Elementary Level

Now that you have completed the past academic year of implementation, please complete this survey to obtain information that will aid in determining the effectiveness and usefulness of Academic Team-initiated Problem Solving for your elementary school. Please read the following statements regarding the primary prevention plan developed by your school and circle the number which best describes your agreement or disagreement with each statement.

	<i>Strongly Disagree</i>	<i>Disagree</i>	<i>Slightly Disagree</i>	<i>Slightly Agree</i>	<i>Agree</i>	<i>Strongly Agree</i>
1. Academic Team-initiated Problem Solving was an acceptable intervention for the elementary school.	1	2	3	4	5	6
2. Most teachers found Academic Team-initiated Problem Solving to be appropriate.	1	2	3	4	5	6
3. Academic Team-initiated Problem Solving was effective in meeting the purposes.	1	2	3	4	5	6
4. I would suggest the use of Academic Team-initiated Problem Solving to other teachers.	1	2	3	4	5	6
5. Academic Team-initiated Problem Solving was appropriate to meet the school's needs and mission.	1	2	3	4	5	6
6. Most teachers found Academic Team-initiated Problem Solving suitable for the described purposes and mission.	1	2	3	4	5	6
7. I used Academic Team-initiated Problem Solving in the school setting.	1	2	3	4	5	6
8. Academic Team-initiated Problem Solving did <i>not</i> result in negative side-effects for the students.	1	2	3	4	5	6
9. Academic Team-initiated Problem Solving was appropriate for a variety of students.	1	2	3	4	5	6
10. Academic Team-initiated Problem Solving was consistent with those I have used in school settings.	1	2	3	4	5	6
11. Academic Team-initiated Problem Solving was a fair way to fulfill the intervention purposes.	1	2	3	4	5	6
12. Academic Team-initiated Problem Solving was a reasonable way to meet the stated purposes.	1	2	3	4	5	6
13. I liked the procedures used in Academic Team-initiated Problem Solving.	1	2	3	4	5	6
14. Academic Team-initiated Problem Solving was a good way to meet the specified purposes.	1	2	3	4	5	6
15. The monitoring procedures were manageable.	1	2	3	4	5	6
16. The monitoring procedures gave the necessary						

APPENDIX E

Adapted Self-Assessment of Contextual Fit

Adapted Self-Assessment of Contextual Fit in Schools

Horner, Salentine, & Albin, 2003

The purpose of this interview is to assess the extent to which the elements of a data-based decision making model fit the contextual features of your school environment. The interview asks you to rate (a) your knowledge of the elements of the model, (b) your perception of the extent to which the elements of the model are consistent with your personal values, and skills, and (c) the school's ability to support implementation of the model. This information will be used to design practical procedures that will help school personnel support children with reading challenges. The information you provide will be maintained and reported in a confidential manner consistent with the standards of the American Psychological Association. You will never be identified.

Please think about your experiences learning and implementing the Academic Team-initiated Problem Solving model of decision-making and provide your perceptions of the model. Thank you for your contribution and assistance.

Name of Interviewee: _____ Role : _____

Knowledge of elements in the Decision-making Model.

1. I am aware of the elements of this decision-making model.

1	2	3	4	5	6
Strongly Disagree	Moderately Disagree	Barely Disagree	Barely Agree	Moderately Agree	Strongly Agree

2. I know what I am expected to do to implement this decision-making model.

1	2	3	4	5	6
Strongly Disagree	Moderately Disagree	Barely Disagree	Barely Agree	Moderately Agree	Strongly Agree

Skills needed to implement the Decision-making Model

3. I have the skills needed to implement this decision-making model.

1	2	3	4	5	6
Strongly Disagree	Moderately Disagree	Barely Disagree	Barely Agree	Moderately Agree	Strongly Agree

REFERENCES CITED

- Algozzine, B., Barrett, S., Eber, L., George, H., Horner, R., Lewis, T., & Sugai, G. (2014). School-wide PBIS tiered fidelity inventory. *OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports*, 1-30.
- Algozzine, B., Newton, J. S., Horner, R. H., Todd, A. W., & Algozzine, K. (2012). Development and technical characteristics of a team decision-making assessment tool: Decision observation, recording, and analysis (DORA). *Journal of Psychoeducational Assessment*, 30(3), 237-249.
- Alonzo, J., Tindal, G., Ulmer, K., & Glasgow, A. (2006). easyCBM online progress monitoring assessment system. *Eugene, OR: Center for Educational Assessment Accountability*.
- Baker, S. K., Smolkowski, K., Smith, J. M., Fien, H., Kame'enui, E. J., & Thomas Beck, C. (2011). The impact of Oregon Reading First on student reading outcomes. *The Elementary School Journal*, 112(2), 307-331.
- Balu, R., Zhu, P., Doolittle, F., Schiller, E., Jenkins, J., & Gersten, R. (2015). Evaluation of Response to Intervention Practices for Elementary School Reading. NCEE 2016-4000. *National Center for Education Evaluation and Regional Assistance*.
- Bandeira de Mello, V., Bohrnstedt, G., Blankenship, C., & Sherman, D. (2015). Mapping State Proficiency Standards onto NAEP Scales: Results from the 2013 NAEP Reading and Mathematics Assessments. NCES 2015-046. *National Center for Education Statistics*.
- Butler, S. R., Marsh, H. W., Sheppard, M. J., & Sheppard, J. L. (1985). Seven-year longitudinal study of the early prediction of reading achievement. *Journal of Educational Psychology*, 77(3), 349.
- Cattell, J. M. (1886). The time it takes to see and name objects. *Mind*, 11(41), 63-65.
- Consortium, S. B. A. (2016). Smarter Balanced Assessment Consortium: 2014-15 Technical Report. In: Los Angeles: Author.
- Coyne, M. D., Kame'enui, E. J., & Simmons, D. C. (2004). Improving Beginning Reading Instruction and Intervention for Students with LD: Reconciling "All" with "Each". *Journal of Learning Disabilities*, 37(3), 231-239.
- D'zurilla, T. J., & Goldfried, M. R. (1971). Problem solving and behavior modification. *Journal of abnormal psychology*, 78(1), 107.
- Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional children*, 52(3), 219-232.

- Filderman, M. J., Toste, J. R., Didion, L. A., Peng, P., & Clemens, N. H. (2018). Data-based decision making in reading interventions: a synthesis and meta-analysis of the effects for struggling readers. *The Journal of Special Education, 52*(3), 174-187.
- Fixsen, D. L., Blase, K. A., Naoom, S. F., & Wallace, F. (2009). Core implementation components. *Research on social work practice, 19*(5), 531-540.
- Fixsen, D. L., Naoom, S. F., Blase, K. A., & Friedman, R. M. (2005). Implementation research: a synthesis of the literature.
- Flesch, R. (1955). Why Johnny can't read: And what you can do about it.
- Foorman, B., Beyler, N., Borradaile, K., Coyne, M., Denton, C. A., Dimino, J., . . . Justice, L. (2016). Foundational Skills to Support Reading for Understanding in Kindergarten through 3rd Grade. Educator's Practice Guide. NCEE 2016-4008. *National Center for Education Evaluation and Regional Assistance*.
- Fuchs, D., & Fuchs, L. S. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading research quarterly, 41*(1), 93-99.
- Fuchs, L. S., & Fuchs, D. (2007). A model for implementing responsiveness to intervention. *Teaching Exceptional Children, 39*(5), 14-20.
- Gersten, R., Beckmann, S., Clarke, B., Foegen, A., Marsh, L., Star, J. R., & Witzel, B. (2009). Assisting Students Struggling with Mathematics: Response to Intervention (RtI) for Elementary and Middle Schools. NCEE 2009-4060. *What Works Clearinghouse*.
- GlobalScholar. (2011). *Response to intervention (RTI) adoption survey 2011*. Retrieved from Bellevue, WA:
- Good III, R. H., & Kaminski, R. A. (1996). Assessment for instructional decisions: Toward a proactive/prevention model of decision-making for early literacy skills. *School Psychology Quarterly, 11*(4), 326.
- Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*: Routledge.
- Horner, R., Salentine, S., & Albin, R. (2003). Self assessment of contextual fit in schools. *Eugene, OR: Educational and Community Supports*.
- Horner, R. H., Newton, J. S., Todd, A. W., Algozzine, B., Algozzine, K., Cusumano, D., & Preston, A. (2018a). A Randomized Waitlist Controlled Analysis of Team-Initiated Problem Solving Professional Development and Use. *Behavioral Disorders, 43*(4), 444-456. doi:10.1177/0198742917745638

- Horner, R. H., Newton, J. S., Todd, A. W., Algozzine, B., Algozzine, K., Cusumano, D., & Preston, A. (2018b). A Randomized Waitlist Controlled Analysis of Team-Initiated Problem Solving Professional Development and Use. *Behavioral Disorders*, 0198742917745638.
- Horner, R. H., & Sugai, G. (2015). School-wide PBIS: An Example of Applied Behavior Analysis Implemented at a Scale of Social Importance. *Behavior Analysis in Practice*, 8(1), 80-85. doi:10.1007/s40617-015-0045-4
- Horner, R. H., Sugai, G., Smolkowski, K., Eber, L., Nakasato, J., Todd, A. W., & Esperanza, J. (2009). A randomized, wait-list controlled effectiveness trial assessing school-wide positive behavior support in elementary schools. *Journal of Positive Behavior Interventions*, 11(3), 133-144.
- Huey, E. B. (1908). *The psychology and pedagogy of reading*: The Macmillan Company.
- Jerrolds, B. W. (1977). Reading Reflections: The History of the International Reading Association.
- Joyce, B., & Showers, B. (1982). The coaching of teaching. *Educational leadership*, 40(1), 4.
- Joyce, B. R., & Showers, B. (2002). Student achievement through staff development.
- Kratochwill, T. R., Hitchcock, J., Horner, R., Levin, J. R., Odom, S., Rindskopf, D., & Shadish, W. (2010). Single-case designs technical documentation. *What Works Clearinghouse*.
- Lane, K., Robertson, E., & Wehby, J. (2002). Primary intervention rating scale. *Unpublished rating scale*.
- May, S., Ard, W., Todd, A., Horner, R., Glasgow, A., Sugai, G., & Sprague, J. (2003). School-wide information system. *Eugene: Educational and Community Supports, University of Oregon*.
- McIntosh, K., & Goodman, S. (2016). *Integrated multi-tiered systems of support: Blending RTI and PBIS*: Guilford Publications.
- McIntosh, K., Massar, M. M., Algozzine, R. F., George, H. P., Horner, R. H., Lewis, T. J., & Swain-Bradway, J. (2017). Technical adequacy of the SWPBIS tiered fidelity inventory. *Journal of Positive Behavior Interventions*, 19(1), 3-13.
- Monzalve, M. (2016). *EXAMINING THE RELATION BETWEEN CONTEXTUAL FIT AND IMPLEMENTATION FIDELITY ON BEHAVIOR SUPPORT PLANS AND STUDENT OUTCOMES*. University of Oregon.

- Newton, J., Todd, A., Algozzine, K., Horner, R., & Algozzine, B. (2009). The team initiated problem solving (TIPS) training manual. *Educational and Community Supports, University of Oregon unpublished training manual.*
- Newton, J. S., Algozzine, B., Algozzine, K., Horner, R. H., & Todd, A. W. (2011). Building local capacity for training and coaching data-based problem solving with positive behavior intervention and support teams. *Journal of Applied School Psychology, 27*(3), 228-245.
- Newton, J. S., Horner, R. H., Algozzine, B., Todd, A. W., & Algozzine, K. (2012a). A randomized wait-list controlled analysis of the implementation integrity of team-initiated problem solving processes. *J Sch Psychol, 50*(4), 421-441.
doi:10.1016/j.jsp.2012.04.002
- Newton, J. S., Horner, R. H., Algozzine, B., Todd, A. W., & Algozzine, K. (2012b). A randomized wait-list controlled analysis of the implementation integrity of team-initiated problem solving processes. *Journal of School Psychology, 50*(4), 421-441.
- Nezu, A. M., Nezu, C. M., & Perri, M. G. (1989). *Problem-solving therapy for depression: Theory, research, and clinical guidelines*: John Wiley & Sons.
- NICHD, N. R. P. (2000). *Report of the national reading panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*: National Institute of Child Health and Human Development, National Institutes of Health.
- Shanahan, T., Callison, K., Carriere, C., Duke, N. K., Pearson, P. D., Schatschneider, C., & Torgesen, J. (2010). Improving Reading Comprehension in Kindergarten through 3rd Grade: IES Practice Guide. NCEE 2010-4038. *What Works Clearinghouse.*
- Sharp, K., Sanders, K., Noltemeyer, A., Hoffman, J., & Boone, W. J. (2016). The Relationship Between RTI Implementation and Reading Achievement: A School-Level Analysis. *Preventing School Failure, 60*(2), 152-160.
doi:10.1080/1045988X.2015.1063038
- Shinn, M., & Shinn, M. (2002). AIMSweb training workbook: Administration and scoring of reading curriculum-based measurement (R-CBM) for use in general outcome measurement. *Eden Prairie, MN: Edformation.*
- Stainthorp, R., & Hughes, D. (2004). What happens to precocious readers' performance by the age of eleven? *Journal of Research in Reading, 27*(4), 357-372.

Stockard, J., Wood, T. W., Coughlin, C., & Rasplica Khoury, C. (2018). The effectiveness of direct instruction curricula: A meta-analysis of a half century of research. *Review of Educational Research, 88*(4), 479-507.

Todd, A. W., Horner, R. H., Newton, J. S., Algozzine, R. F., Algozzine, K. M., & Frank, J. L. (2011). Effects of team-initiated problem solving on decision making by schoolwide behavior support teams. *Journal of Applied School Psychology, 27*(1), 42-59.