FUNCTIONAL BEHAVIORAL ASSESSMENT OF PROBLEM-SOLVING TEAMS: ADAPTATION OF THE PERFORMANCE DIAGNOSTIC CHECKLIST FOR APPLICATION IN SCHOOLS

By

Stephanie Sorensen

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Educational Psychology)

at the

UNIVERSITY OF WISCONSIN-MADISON

2015

Date of final oral examination: 4/28/2015

This dissertation is approved by the following members of the Final Oral Committee:
Jennifer Asmus, Professor, Educational Psychology
Thomas Kratochwill, Professor, Educational Psychology
Stephen Quintana, Professor, Counseling Psychology and Educational Psychology
James Wollack, Professor, Educational Psychology
Kimber Wilkerson, Professor, Rehabilitation Psychology and Special Education

Acknowledgements

Firstly, I would like to thank the faculty of the UW-Madison School Psychology Program. You have each contributed to my professional development in your own unique way. Tom Kratochwill and Jennifer Asmus worked tirelessly assisting me with this dissertation project. Steve Quintana has been unwavering in his support – both personally and professionally – as my academic advisor. Julia McGivern is an impassioned clinical supervisor, without whom I would not be half the clinician I am today. Maribeth Gettinger has been an inspiration for what a teacher, leader, and a woman can be. And Craig Albers models an enthusiasm for his work that is contagious.

Thank you to the remainder of my dissertation committee: James Wollack, Kimber Wilkerson, and Elizabeth Dohrn. Your feedback and our discussions pushed me to think critically about this dissertation study and its implications for the field. Jim, I would not have been able to complete this project without your statistical expertise and assistance.

I would also like to thank all of the amazing individuals I had the opportunity to navigate graduate life with. Without your countless study dates, endless moral support, and occasional spontaneous adventures I would likely be a recipient of intensive mental health services rather than a provider of them. You all taught me just as much about our field than any text or class ever could have.

I have been incredibly fortunate to have received clinical and field supervision from the most highly qualified and talented group of professionals: Mya Collins, Kristen Guetschow, Julia McGivern, Robert Eldredge, Mark Hedges, Jennifer Burkett, and Ted Williams, you have all demonstrated levels of excellence that I will spend my career aspiring to achieve. Thank you for everything you have done for me. It has made all the difference in the world.

My new work family at the South Dakota Developmental Center deserves thanks as well, for being so supportive of me during this process. A special thanks to Ted Williams, Darren Jilek, and Deb Sanger for their listening ears, understanding, and ongoing support.

Lastly and most importantly, I would like to thank my family and best friends. You have all been absolutely relentless in your support of my academic and professional endeavors. None of this would mean a thing without you all to share my struggles and accomplishments with. I am forever grateful.

Mom and Dad – thank you for believing in me when I didn't believe in myself, and for pushing me when I felt that I had nothing left to give. I owe you everything.

Table of Contents

List of Tables

List of Figures

Abstract of Dissertation Presented to the Graduate School of the
University of Wisconsin-Madison in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

By
Stephanie Sorensen, M.S.
April 2015
Chairs: Jennifer Asmus, Ph.D. and Thomas Kratochwill, Ph.D.
Committee Members: Stephen Quintana, Ph.D., James Wollack, Ph.D.,
and Kimber Wilkerson, Ph.D.
Major Department: Educational Psychology

Despite widespread regulations, callings, and efforts demanding the increased implementation of evidence-based interventions (EBIs) within the school setting, EBIs are not being readily adopted for students experiencing social, emotional, and/or behavioral challenges. School-based problem-solving teams (PSTs) are being increasingly looked to as the units responsible for selecting and implementing EBIs, yet many barriers exist preventing PSTs from doing so. However, specifically what these barriers are is often unknown or overlooked. In order to create interventions aimed at increasing PSTs' uptake of EBIs, an assessment protocol needs to be developed that can help to identify the barriers experienced by teams.

The current study involved an adaptation of the *Performance Diagnostic Checklist* (PDC), created for application in the Organizational Behavior Management (OBM) field, into a checklist relevant for school PSTs (the *Performance Diagnostic Checklist for Schools*; PDC-S). The PDC-S is grounded in Applied Behavior Analysis and Functional Behavioral Assessment/Analysis, and yields functional assessment information to inform the development of function-based interventions to improve PSTs' uptake of EBIs. In addition to developing the PDC-S, this study examined its content validity, inter-rater reliability, and test-retest reliability.

Results suggest that the psychometric properties of the PDC-S are inconsistent. While the Content Validity Indices were high for the measure, the Factorial Validity Indices did not

meet standards. Cronbach alphas were strong between administration one and administration two, indicating strong test-retest reliability. Percent agreement between PST members was moderate. However, there was little difference between within-group ratings and between-group ratings, suggesting that at least a portion of the percent agreement found between PST members was due to chance. Thus, the inter-rater reliability of the PDC-S was not strong.

CHAPTER 1

Introduction

The increasing prevalence of childhood mental, emotional, and behavioral (MEB) disorders has called attention to the role that schools should play in addressing these disorders. Approximately 14% to 20% of children experience a MEB disorder, and the influence of these disorders on children's lives is significant (National Research Council and Institute of Medicine [NRCIM], 2009). Despite the widespread effects MEB disorders have on youth and their social systems, only 20% to 25% of children with MEB disorders receive treatment (Masi & Cooper, 2006).

Recent research and legislation, such as No Child Left Behind (NCLB) and Individuals with Disabilities Education Act (IDEA), have called to attention the need for evidence-based interventions (EBIs) to treat the MEB disorders experienced by today's youth (IDEA, 2004; NCLB, 2002). Specifically, schools are being called upon to increase their use of EBIs to both prevent the onset of MEB disorders and intervene with students facing mental, emotional, and behavioral challenges. Many schools are utilizing problem-solving teams (PSTs) to address students coping with MEB disorders. However, the rate or ease at which PSTs are selecting and implementing EBIs could be increased to meet need (DuPaul, 2003; Kratochwill, 2007; Schaughency & Ervin, 2006; Walker, 2004). Yet, there is currently little being done to address the low uptake of EBIs by PSTs.

The Institute of Education Sciences (IES) recently funded a grant proposal (R324A12021) entitled *Systems-Level Analysis of Evidence-Based Intervention Implementation by Problem-Solving Teams,* by Dr. Thomas Kratochwill and Dr. Jennifer Asmus through the Wisconsin Center of Educational Research (WCER) at the University of Wisconsin-Madison.

The goal of this project is to apply principles of functional behavioral analysis (FBA) to school problem-solving teams (PSTs) to examine barriers associated with the teams' selection and implementation of EBIs. A critical assessment tool identified for gathering these data will be an adaptation of the *Performance Diagnostic Checklist* (PDC), which was created for use within the Organizational Business Management (OBM) field (Austin, Carr, & Agnew, 1999). The purpose of this study is to adapt the PDC for use with PSTs. Information gathered via the PDC will then be used to inform the development of team-level interventions to increase PSTs uptake of EBIs.

CHAPTER 2

Review of Literature

Behavioral Disorders in Youth

Approximately 6% of children experience a disruptive behavior disorder **(**Costello, Mustillo, Keeler, & Angold, 2004). Behavior disorders can have serious consequences in a child's life, affecting home life, peer relationships, and academic performance (Zins, Bloodworth, Weissberg, & Walberg, 2004). Additionally, behavior disorders affect these children's parents, siblings, and the schools they attend. Unaddressed behavior disorders can be disruptive to classrooms and schools, affecting teachers and other students as well (Reinke, Stormont, Herman, Puri, & Goel, 2011). The prevalence of children experiencing behavior disorders, and the subsequent implications for them and their social systems, call for effective interventions to help protect against the deleterious effects of behavior disorders. However, despite the high rates of MEB disorders among children, 75% to 80% of children's mental health needs are not met (Masi & Cooper, 2006). Of the 20% to 25% of children with MEB disorders receiving mental health services, the vast majority are receiving these services within an educational setting (i.e., school; Farmer, Burns, Phillips, & Costello, 2003; Hoagwood, Burns, Kiser, Ringeisen, & Schoenwald, 2001; Walker, 2004).

Evidence-Based Practices

In response to the high rates of behavior disorders in children, and low treatment of these disorders, many organizations, researchers, advocates, and even legislation have called for an emphasis on the application of evidence-based practices (EBPs) in community and school settings (IDEA, 2004; Kratochwill & Stoiber, 2000; NCLB, 2002; Psychological Association Task Force on Evidence-Based Practice for Children and Adolescents [APA Task Force], 2008).

The No Child Left Behind Act of 2002 requires schools to use practices that have been scientifically validated. The Individuals with Disabilities Education Act (IDEA, 2004) also require that schools implement scientifically based practices when working with students.

While terminology may differ across sources, the meaning of the terms *scientifically validated practices, evidence-based practices, empirically-based practices, and research-based practices* are similar. The APA Task Force (2008) presented the following definition for EBP in psychology: "Evidence-based practice in psychology (EBPP) is the integration of the best available research with clinical expertise in the context of patient characteristics, culture, and preferences" (p. 6). Although sometimes used interchangeably, it is important to clarify the differences between practices and/or interventions identified as being efficacious and those proven to be effective. An intervention identified as being "efficacious" means that it has yielded significant, positive outcomes in controlled contexts, typically in randomized treatment studies. However, effectiveness refers to interventions that have produced significant, positive outcomes in the settings for which they were intended (Schoenwald & Hoagwood, 2001).

To facilitate the implementation of EBPs in schools, or what will hereafter be referred to as evidence-based interventions (EBIs), several registries have been created for use by school personnel to search for EBIs. Examples of these registries include the Institute of Education Science's (IES's) What Works Clearinghouse (WWC) and Substance Abuse and Mental Health Services Administration (SAMSHA).

Research-practice gap in EBI implementation. Despite recognition of the need for EBIs, legislation requiring the use of EBIs, and steps to make EBIs more available to school personnel, schools have failed to implement EBIs to their full capacity (DuPaul, 2003; Kratochwill, 2007; Schaughency & Ervin, 2006; Walker, 2004). The literature identifies many

barriers that prevent schools from selecting and implementing EBIs, including knowledge and training; time and resources; and school personnel resistance (Conoley, Conoley, & Reese, 2009; Reinke et al., 2011). While the need for EBIs for children with behavior disorders continues to increase, teachers' feelings of competency in this area have not. Many teachers report feeling overwhelmed by the prospects of managing students with challenging behaviors (Lopes, Monteiro, & Sil, 2004). One way that this issue has been and is being addressed is through the formation of school problem-solving teams (PSTs).

Problem-Solving Teams

Although the implementation of PSTs has received significant attention in recent years, it is not a new concept to the field of education (Burns, Wiley, & Viglietta, 2008b; Graden, Casey, & Christenson, 1985). In fact, school personnel have been collaborating to improve student outcomes for decades. Whether called teacher assistance teams (TATs; Chaflant, VanDusen Pysh, & Moultrie, 1979), prereferral intervention teams (PITs; Graden, et al., 1985), or instructional consultation teams (ICTs; Rosenfield & Gravois, 1996), the concept of collaborating within school buildings is a long-standing tradition in the field of education. However, the recent increase in school PSTs can, in large part, be linked to legislative mandates including IDEA (2004) and NCLB (2002; Prasse, 2006). Specifically, IDEA (2004) allows schools to use up to 15% of their special education funding to provide preventative and intervention services to students not receiving special education services. Prasse (2006) states, "This provision clearly supports functional problem-solving approaches, such as noncategorical services, and the implementation of problem-solving strategies for the coordination and integration of education, health, mental health, and social services" (p. 9).

PST members may vary from district to district; however, most teams include the teacher with a student referral, a supporting teacher, the principal, a special education teacher, and at least one additional specialist (e.g., school psychologist, school counselor, and/or social worker; Burns et al., 2008b; Iverson, 2002). Although the inclusion of special education teachers and school psychologists have been controversial given PSTs' focus on students in general education, research shows that teams which included these members were more effective (Burns, 1999). To avoid what some scholars might refer to as "admiration of the problem," it is recommended that PSTs spend approximately 10 to 15 minutes per child (Allen & Graden, 2002; Burns et al., 2008a). The role of PSTs is to assist teachers with implementing EBIs to improve student outcomes, with the goal of reducing the number of student referrals for special education (Bahr & Kovaleski, 2006). Thus, PSTs can be thought of as a form of problem-solving consultation for teachers where they engage in a systematic problem-solving process (Kratochwill, 2008).

Although PSTs vary widely in their processes across schools and districts, there are some components of PSTs that are consistent across the majority of teams. Specifically, the steps taken when a student referral is made typically follow the same sequence: (1) teacher requests consultation; (2) consultation occurs between teacher and other school personnel, wherein an intervention is selected and implemented; (3) if progress monitoring data shows that the intervention is not effective, additional data is collected via observations (e.g., a functional behavior assessment); (4) the team meets to make a decision based on the additional data; and (5) if necessary, a formal referral for special education is made (Burns & Symington, 2002; Graden et al., 1985).

Beyond these basic steps, teams typically operate one of two ways within their team processes: a problem-solving approach or a standard protocol approach (Fuchs & Fuchs, 2006; Fuchs, Mock, Morgan, & Young, 2003). A problem-solving approach involves selection and implementation of interventions based on the referred student's individual problem behavior, as identified by a comprehensive, individualized assessment (Fuchs & Fuchs, 2006). Conversely, a standard protocol approach involves the implementation of a standard battery of interventions for each child referred (Jimerson, Burns, & VanDerHeyden, 2007). In the latter approach, the PST applies the same intervention for each student referred to the PST for similar types of behavioral concerns. While support exists for the use of both models, best practice recommends utilizing a problem-solving approach (Tilly, 2008). Additionally, NCLB and IDEA require the use of problem-solving processes prior to referring a student for special education (IDEA, 2004; NCLB, 2002; Prasse, 2006).

The problem-solving process. The problem-solving approach follows four procedural steps in assessing and intervening with children referred to the PST by their teacher: (1) problem identification, (2) problem analysis, (3) treatment implementation, and (4) treatment evaluation (Sheridan & Kratochwill, 2008; Tilly, 2008). This process should not be viewed as linear, but rather cyclical; when a treatment is not effective in addressing the problem, the process should be repeated. During the first step of problem identification, the team works to objectively define the problem behavior(s) exhibited by the student. In the problem analysis phase, the team collects assessment data to determine why the problem behavior is occurring. Once hypotheses have been generated regarding the reason for the problem, the team selects a treatment plan and begins implementation. During the fourth step, treatment evaluation, two evaluations should occur: (1) the integrity with which the treatment plan, or intervention, is being implemented, and (2) the

students' response to the intervention (Sheridan & Kratochwill, 2008; Tilly, 2008). When this process is implemented with integrity, positive student outcomes have been observed (Burns & Symington, 2002; Doll, Gaack, Kosse, Osterloh, & Siemers, 2005). Specifically, Burns and Symington (2002) conducted a meta-analysis of both system- and student-level outcomes influenced by PSTs and found that student outcomes including task completion, readings skills, and time on task improved and systems issues such as referrals and placements in special education decreased.

Furthermore, school teams which do not follow a structured, problem-solving process tend to see fewer positive student outcomes and have lower teacher approval of the referral team process. Specifically, general education teachers involved with such teams reported feeling that there was insufficient time to discuss and address student concerns; a lack of focus during team meetings; and disagreement among team members regarding the nature of students' referral concerns (Williamson & McLeskey, 2011). Disagreement can also exist among team members regarding the importance of the consultation process (Nellis, 2012).

Evaluation of PSTs. Various measures have been developed to assess the overall functioning of PSTs and/or PST processes. For example, the *Decision Observation, Recording, and Analysis, Second Edition* (DORA II; Algozzine et al., 2011) is a direct observation tool that provides information regarding team processes. Specifically, the DORA II examines whether PSTs establish outcome goals; use data; develop intervention(s); devise progress monitoring plans; establish methods of monitoring implementation fidelity; and follow-up on referred students' progress (Algozzine et al., 2011). In addition to this direct measure, indirect measures such as the *Problem-Solving Team Process Fidelity Checklist* (Burns et al., 2008a) and *The Team Functioning Scale* (Gaumer Erickson et al., 2014) evaluate other elements of PSTs and the PST

processes. However, there are no measures – direct or indirect – currently available to collect information regarding why a PST is not selecting and implementing EBIs.

Problem-solving teams' use of EBIs. As discussed above, NCLB and IDEA mandate the use of scientifically based research, or EBIs, to address student behavior (IDEA, 2004; NCLB, 2002). While the development and availability of EBIs is increasing, they are still not being readily adopted and implemented in school settings (DuPaul, 2003; Kratochwill, 2007; Schaughency & Ervin, 2006). The literature is relatively silent regarding reasons for this research-to-practice gap. However, teachers' perceptions regarding their own knowledge and skills, as well as their views of consultation, have been identified as contributing factors to the gap (Lopes et al., 2004; Reinke et al., 2011). In particular, Williamson and McLeskey (2011) found that some teachers feel that PST meetings are nothing more than a time for teachers to be critically evaluated, and sometimes blamed for students' challenges and/or lack of success.

Teachers are an integral part of PSTs and the problem-solving process. Teachers' responsibilities may include identifying children in need of behavioral interventions and supports; referring students to the PST; implementing interventions; and monitoring students' response to the implemented intervention (Reinke et al., 2011). Hence, it is crucial that teachers' have a sound knowledge base regarding MEB disorders and EBIs, and feel comfortable implementing behavioral interventions. Unfortunately, the majority of teachers feel they lack the training and skills necessary to effectively intervene with students with behavioral disorders (Lopes et al., 2004; Reinke et al., 2011). Another barrier to selecting and implementing EBIs in school settings via PSTs are teachers' perceptions of consultation, including its efficacy and their own roles and responsibilities within the process (Conoley et al., 2009; Nellis, 2012). The

failure to fit with a school's primary mission can also be a barrier in the adoption of EBIs (Adelman & Taylor, 2003; Walker, 2004).

It is important to note that the implementation of EBIs to address childhood behavioral disorders is not unique to schools or school-based PSTs. This research-to-practice gap exists in other fields including community health settings (APA Task Force, 2008). In particular, Glisson and colleagues (Glisson et al., 2008; Glisson & Schoenwald, 2005) have recognized the need for organizational change to improve the implementation of effective interventions for MEB disorders in community settings. This focus includes the development of methods to assess the social context of organizations, as well as the identification of factors within the organizations that influence the implementation of EBPs (Glisson, 2007). While the work of Glisson and colleagues is related to the goals of the current IES grant within which this dissertation project resides, there are two major differences: (a) the current grant and study are specifically focusing on the selection and implementation of EBIs by school PSTs, rather than community mental health agencies/settings, and (b) this grant and the current study are utilize principles and procedures of Applied Behavior Analysis (ABA) and Functional Behavioral Assessment/Analysis (FBA) to conduct assessments and gather information (Kratochwill & Asmus, 2011).

IES Grant

Given the prevalence and implications of behavior disorders among youth, as well as the lack of school PSTs' selection and implementation of EBIs to address these disorders, IES funded a grant proposal by Kratochwill and Asmus (R324A12012) specifically addressing the uptake of EBIs by PSTs. This grant is officially entitled *Systems-Level Analysis of Evidence-Based Intervention Implementation by Problem-Solving Teams,* but is more frequently referred to

as the *Learning to Improve School Teams* (LIST) Project. For the purposes of this paper, the names LIST, or LIST Project, will be used to refer to the grant.

The primary goal of the LIST Project is "…to develop, implement, and empirically evaluate an assessment and intervention protocol designed to increase PST's adoption and implementation of EBPs targeted at students with disruptive behavior problems and disorders" (Kratochwill & Asmus, 2008, p.1). The grant proposal was to conduct FBAs of PSTs and generate testable hypotheses regarding possible barriers to the selection and implementation of EBIs by each team. By applying methods of FBA, LIST researchers will be able to design and implement individualized interventions with the PSTs to increase their selection and implementation of EBIs. The next section of this paper will discuss principles and procedures of both Applied Behavior Analysis (ABA) and FBA.

Applied Behavior Analysis

In order to discuss ABA, one must first review the important works and contributions of B.F. Skinner. Skinner (1953) emphasized the influence environmental factors have on behavior, including their ability to increase, decrease, or maintain the frequency a behavior occurs, also known as operant conditioning or operant behaviorism. This influence has been referred to as the functional relationship between conditions and behavior (Alberto & Troutman, 2009). As will be discussed later, identifying the function, or purpose, of a behavior is paramount to the FBA process (O'Neill et al., 1997). However, Skinner's work primarily studied the behavior of animals, namely white rats and pigeons.

It wasn't until the 1960's when behaviorists began applying these principles to humans in their natural environments (Alberto & Troutman, 2009; Baer, Wolf, & Risley, 1968; Bijou, Peterson, & Ault, 1968; Alberto & Troutman, 2009). In their seminal article, Baer, Wolf, and

Risley (1968) defined ABA as the "process of applying sometime tentative principles of behavior to the improvement of specific behaviors, and simultaneously evaluating whether or not any changes noted are indeed attributed to the process of application" (p. 91). The main tenants of ABA include a focus on behaviors considered to be socially important, and studying behaviors in their naturally occurring setting rather than laboratories (Baer et al., 1968). Carr (1977) later expanded on this theory and proposed three hypotheses for the motivation of reoccurring problematic behavior: positive reinforcement, negative reinforcement, and sensory reinforcement.

Functional Behavioral Analysis

Steege and Pratt (2013) describe FBA as "…both (1) a theoretical framework for understanding human behavior and (2) a set of assessment procedures" (p. 126). As a theory, FBA highlights the interactions between individuals and the environments within which they operate (Steege & Pratt, 2013; Steege & Watson, 2009). As a set of assessment procedures, the primary goal of FBA is to conduct a multimodal evaluation of a child's challenging behavior(s) to identify the function, or purpose, for that behavior (Gresham, Watson, & Skinner, 2001). Hanley, Iwata, and McCord (2003) present two uses of the term function: (1) "…the effect that a behavior has on the environment, or speaking loosely, the purpose the behavior serves for an individual (e.g., the function of the behavior is to terminate an ongoing event)", and (2) "…a relationship between two variables (typically between some environmental event and a class of behavior) in which one varies given the presence or absence of the other (e.g., responding as a function of an event)." Identifying the function of a behavior is important because it allows for the development of an intervention that directly replaces the challenging or negative behavior with a socially appropriate behavior that serves the same function (Steege & Pratt, 2013).

Furthermore, when an intervention is developed based on the function of a problem behavior, it is more effective than an intervention based on a diagnosis (Betz & Fisher, 2011).

There is a process to conducting FBAs that, when followed, yields testable hypotheses regarding the setting events, antecedents, and consequences maintaining the problem behavior (O'Neill et al., 1997; Sugai, Lewis-Palmer, & Hagan-Burke, 2000). The development of these testable hypotheses, also known as summary statements, is crucial as it increases the relevancy and effectiveness of subsequent function-based interventions (O'Neill et al.; Sugai et al., 2000). To collect all the necessary information for a FBA, a team should be assembled (Asmus, Vollmer, & Correro, 2002). Team members may include family members, teachers, support staff, and, when possible, students (O'Neill et al., Sugai et al., 2000). There are three major methods utilized during a FBA to collect necessary information: indirect, direct, and experimental or functional analysis (Gresham et al., 2001; O'Neill et al., 1997; Steege & Watson, 2009). Indirect and direct methods of FBA assessments are both considered descriptive (Bijou et al., 1968). Indirect methods include the gathering of information from various informants closest to the client via interviews, rating scales, and/or checklists. These methods of gathering data can aid in creating a description of the client's behavior patterns, identifying any variables that may influence the problem behavior, determining which environmental settings act as triggers for the problem behavior (O'Neill et al., 1997). Direct methods involve observations in the setting where the behavior(s) naturally occur. Antecedent-behavior-consequence (A-B-C) recordings are commonly used as a form of direct methods. A-B-C recordings allow the team to observe when a problem behavior occurs (behavior) events that took place just prior to the behavior (antecedent), events that took place immediately following the behavior (consequence), and hypotheses regarding the function of the behavior (O'Neill et al., 1997). A FBA requires the

manipulation of variables to determine the true function of the challenging behavior (Gresham et al., 2001; Iwata, Pace, Cowdery, & Miltenberger, 1994; Steege & Watson, 2009). While best practices involve the use of both descriptive and experimental methods of FBA, the current discussion will focus on indirect, descriptive methods for the purposes of this study (Asmus et al., 2002).

Indirect methods of FBA. FBAs have been used widely in education and psychology at the individual client level. Hence, many methods of descriptive, indirect FBAs have emerged over the years, including the *Functional Assessment Interview* (FAI; O'Neill et al., 1997) and the *Motivation Assessment Scale* (MAS; Durand & Crimmons, 1988). While both of these measures provide useful information for creating function-based interventions, the completion time for the FAI is quite long, and little or no psychometric properties exist supporting these tools' reliability and validity (Matson, Tureck, & Rieske, 2012; Paclawskyj et al., 2000). In response to these shortcomings, Matson and Vollmer (1995) developed the *Questions About Behavioral Function* (QABF), which seeks to gather the same information as the FAI and MAS, but in an abbreviated, checklist form. Additionally, several studies have been conducted supporting the psychometric qualities of the QABF (Matson, Bamburg, & Cherry, 1999; Paclawskyj et al., 2000).

Application of FBA principles to systems. Just as IDEA (2004) requires individualized education program (IEP) teams to conduct functional behavioral analyses for students experiencing emotional and behavioral difficulties as the primary assessment for developing a behavioral intervention, the LIST grant proposes to use an apparatus of tools and measures to collectively perform the assessment goals of an FBA for a PST. That is, researchers will conduct a FBA, including the Performance Diagnostic Checklist (PDC), to identify barriers PSTs experience in selecting and implementing EBIs, and develop interventions to reduce identified

barriers (Kratochwill & Asmus, 2011). While this is unprecedented in the education and psychology fields, areas such as organizational behavior management (OBM) have begun applying FBA principles and procedures to the functioning of organizations, as can be seen in the development and application of the *Performance Diagnostic Checklist* (PDC; Austin, 2000).

Performance Diagnostic Checklist

Austin, Carr, and Agnew (1999) developed the PDC (see Appendix A) for use in organizational business management (OBM). The checklist was based on principles of ABA and FBA to serve as an assessment of antecedents; equipment/materials and processes; knowledge and skills; and consequences that influence employee adherence to job duties and responsibilities. The goal of the PDC was to identify barriers preventing employees from performing their duties with integrity, and then develop interventions that directly address identified barriers (Austin, Carr, & Agnew, 1999). Some studies have found success in using the PDC as an assessment tool for developing function-based interventions aimed at improving employee performance (e.g., Pampino, Heering, Wilder, Barton, & Burton, 2003; Rodriguez et al., 2006; Rohn, Austin, & Lutrey, 2003).

In one study, Rohn et al. (2003) used the PDC to analyze variables resulting in cash register shortages at a business. Researchers gathered PDC information via observations and interviews with employees. Results indicated that problematic equipment and processes, as well as a lack of consequences, maintained the occurrences of cash register shortages. Based on these results, an intervention package consisting of feedback and accountability was designed and implemented. Results from an ABAB reversal design showed significant decreases in the amount of cash shortages during intervention phases, and an increase in cash shortages when the intervention was withdrawn.

In another study involving an independently owned coffee shop, Pampino and colleagues (2003) interviewed five employees using the PDC to identify which of the four major areas of analysis presented the most barriers to maximally performing their closing duties, including restocking and cleaning tasks. The results of the PDC identified antecedents and consequences as the primary areas in need of intervention. Thus, the experimenters developed a packaged intervention addressing both of these areas. Results were examined using a multiple baseline single-case design, and found that the percentage of closing duties completed increased significantly post-intervention.

In a similar study, Rodriguez and colleagues (2006) utilized the PDC to identify variables influencing employees' offerings of promotional stamps at two restaurants. The PDC was administered in interview form with one manager and one employee from each restaurant. PDC results indicated that the following areas exhibited barriers to employees' offerings of promotional stamps: antecedents; equipment and processes; and consequences. Hence, researchers developed and implemented an intervention package including task clarification, employee self-monitoring, equipment modification, goal setting, and graphic feedback. In a multiple baseline design, post-intervention observations demonstrated a 46% and 70% increase in the rate which employees offered customers promotional stamps.

While the PDC was originally developed more specifically for businesses, Carr and colleagues (2013) demonstrated how the PDC can be adapted for application in other contexts (PDC-Human Services, see Appendix B), specifically in an outpatient clinic for individuals with mental health disorders. With some adaptations, the PDC could be a useful tool for schools and PSTs in particular. Schools are, after all, an organization or system quite similar to a business in that they both aim to produce a specific set of outcomes (e.g., high student achievement vs. high

sales) via the adequate performance of trained professionals (e.g., teachers vs. sales associates) performing a specific, unique set of skills (e.g., quality instruction vs. effective sales techniques). Thus, the positive results yielded in the studies above could translate to PSTs. Before using the PDC with school PSTs, the tool will need to be revised and validated by researchers. The current wording of the PDC and PDC-HS do not lend themselves to this purpose. For example, questions such as, "Is there a written description stating exactly what is expected of the employee?" needs to be revised and adapted for a PST in a school-based setting. Additionally, the PDC and PDC-HS have not been evaluated for psychometric properties. Thus, the goal of the current study was to adapt the PDC into the PDC for Schools (PDC-S) to fit the context of school PSTs for the purpose of serving as an indirect FBA tool to aid in the (a) identification of testable hypotheses regarding systems-level barriers to PSTs selecting and implementing EBIs, toward (b) the development of individualized interventions to increase PSTs' uptake of EBIs.

Scale Development

The process of developing a scale can be very intensive. Several scholars have suggested specific guidelines for doing so (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999 [AERA et al., 1999]; DeVellis, 2012; Netemeyer, Bearden & Sharma, 2003), all of which follow similar general steps. DeVellis (2012) proposes the following guidelines for developing scales: (1) determine clearly what you want to measure, (2) generate an item pool, (3) determine the format for measure, (4) have initial item pool reviewed by experts, (5) consider inclusion of validation items, (6) administer items to a development sample, (7) evaluate the items, and (8) optimize scale length. In comparison, Netemeyer and colleagues (2003) recommend the following process: (1) construct definitions and content domain, (2) generate and judge measurement items,

(3) design and conduct studies to develop and refine the scale, and (4) finalize the scale. In essence, Netemeyer and colleagues (2003) have condensed the process proposed by DeVellis (2012). For the purposes of the current study, the process outlined by DeVellis will be discussed and adhered to due to its thoroughness. However, Netemeyer and colleagues' (2003) recommendations will also be explored in the context of DeVellis' (2012) process.

Step 1: Determine clearly what it is you want to measure. One of the most critical, yet often unappreciated, steps in the scale development process is defining precisely what you intend on measuring. Theory plays a significant role in the clarification process for defining the constructs to be measured by a scale. Hence, scholars (DeVellis, 2012; Netemeyer et al., 2003) recommend conducting a thorough literature review in order to fully understand the theory that guides your scale and investigate what, if any, measures already exist that measure the same or similar constructs. By garnering an in-depth understanding of the theory via a thorough review of the literature, scale developers are able to clearly define the construct they intend to measure (Netemeyer et al., 2003).

Step 2: Generate an item pool. Once the construct has been defined, the process of creating items that load onto the construct can begin. The initial item pool generated is typically quite large in number, as the latter steps of the scale development process are intended to eliminate several of the initial items. Thus, it is recommended that scale developers be more inclusive, rather than selective, in their creation and selection of items to add to the item pool (DeVellis, 2012; Netemeyer et al., 2003). There are numerous other factors to consider during the item generation stage of developing a scale, including the wording, clarity, and length of items. Avoiding items that are excessively long is important, as is ensuring that items are written as clearly as possible (DeVellis, 2012).

Step 3: Determine the format for measure. While developers are generating an item pool, consideration should be given to the format that the measure will take. The two major formats scales typically take are dichotomous and multidichotomous. Dichotomous scales typically provide informants with yes-no or true-false response options, whereas multidichotomous scales provide informants with three or more response options typically utilizing Likert scales (Netemeyer et al., 2003).

Step 4: Have initial item pool reviewed by experts. There are many reasons to enlist a panel of experts to review your initial scale, but the most important purpose this step serves is to strengthen a scale's content validity (DeVellis, 2012; Netemeyer et al., 2003). It is recommended that at least five experts judge generated items (Netemeyer et al., 2003). DeVellis (2012) recommends supplying your panel of expert reviewers with definitions of the constructs your scale intends to measure, then asking the reviewers to assign each item from your scale to the most appropriate measure. Reviewers can also be helpful in identifying ambiguous items and suggesting alternative phrasing or wording of items (DeVellis, 2012). A more in-depth description of this process will be discussed later within the context of content validity.

Step 5: Consider inclusion of validation items. After steps one through four have been completed, and the scale is being revised for administration to a developmental sample in step six, developers may consider the inclusion of certain validation items. Scales are susceptible to biased responses from informants, particularly in the form of socially desirable responding (SDR; DeVellis, 2012; Netemeyer et al., 2003). Netemeyer and colleagues (2003) suggest several ways for developers to lessen the likelihood of SDR, including forced-choice items, indirect questioning, randomized response techniques, self-administration of scales, and

anonymity. For the purposes of the current study, self-administration of scales and anonymity are the most appropriate methods that will be used to reduce SDR.

Step 6: Administer items to a development sample. Conducting a pilot administration of the developed scale to a sample is a critical step in the scale development process. There are varied guidelines for determining the appropriate size of samples (*N*), but there is consensus around $N = 300$ being a sufficient pilot sample size (DeVellis, 2012; Netemeyer et al., 2003). However, if a scale has fewer than twenty items and its construct is narrowly defined, $N = 100$ to 200 is considered sufficient (Netemeyer et al., 2003). In addition to recruiting a sufficiently large sample to administer the scale to, it is also important that developers recruit a sample that is representative of the larger population (DeVellis, 2012; Netemeyer et al., 2003).

Step 7: Evaluate the items. After developers have generated, judged, and piloted items "…to an appropriately large and representative sample, it is time to evaluate the performance of the individual items so that appropriate ones can be identified to constitute the scale" (DeVellis, 2012, p. 104). Three major factors should be assessed when evaluating scale items: item-scale correlations, item variances, and item means (DeVellis, 2012). Item-scale correlations, or item-to-total correlations, indicate how well items correlate with the remaining items of a scale (DeVellis, 2012; Netemeyer et al., 2003). DeVellis (2012) identifies relatively high item variances as an important attribute of scales. That is, the sample administration of a scale should yield a large range of scores (i.e., have high item variance) if it is a truly diverse sample. Although the range should be large, item means closest to the middle of the range are most desirable (DeVellis, 2012).

Step 8: Optimize scale length. After evaluating the scale's items during the previous step, developers can use that information to cut items that do not meet recommended standards in

scale development. For example, items with low variance and/or items with skewed means may need to be cut from the pool of items to be incorporated into the final version of the measure (DeVellis, 2012).

Reliability

When developing a scale or measure, it is crucial to ensure that the scale is reliable (DeVellis, 2012). Reliability is concerned with the error of measurement that exists with any measure (AERA, 1999; Nunnally & Bernstein, 1994). The error of measurement, also referred to as *measurement error*, is the difference between the hypothetical true score of a measure and the observed score (AERA, 1999). There are several different types of reliability, including internal consistency, test-retest reliability, and alternative form reliability (DeVellis, 2012; Netemeyer et al., 2003). While these are all important in scale development, the current proposal will focus on test-retest reliability and inter-rater agreement, or inter-rater reliability, as these are the most appropriate types of reliability to analyze for the purposes of the proposed study; "When subjective judgment enters into test scoring, evidence should be provided on…within-examinee consistency over repeated measurements" (AERA et al., 1999, p. 33).

Test-retest reliability. Test-retest reliability can be thought of as the stability of a measure over time (Nunnaly & Bernstein, 1994). That is, scores on an item stay the same, unless there is a change in the variable being measured (DeVellis, 2012). Test-retest reliability is assessed by comparing test scores at administration one and administration two, thus yielding a test-retest reliability coefficient, or the coefficient of stability (Crocker & Algina, 1986; DeVellis, 2012; Webb, Shavelson, & Haertel, 2006). It is recommended that all reliability coefficients be greater than or equal to 0.80 (Webb et al., 2006). There is no concrete rule regarding how much time should elapse between the two administrations. However, it is

recommended that enough time pass so that raters do not remember their responses, but not so much time that a change in responses could reflect an actual change in the construct being measured (Crocker & Algina, 1986). Many scholars suggest that a two-week period between administrations is an appropriate amount of time (Netemeyer et al. 2003; Webb et al., 2006).

Inter-rater reliability. Whenever there are multiple raters being asked to complete the same measure, it is important to examine inter-rater reliability (Crocker & Algina, 1986; DeVellis, 2005/2012). DeVellis (2005) defines inter-rater reliability as, "…the extent to which assessments of a phenomenon by two or more observers are influenced by some aspect of the phenomenon being observed rather than by some aspect of the observers" (p. 317). That is, inter-rater reliability is interested in the agreement across raters. There are several methods available for calculating and assessing the inter-rater reliability, or agreement, of a scale. The simplest method is calculating the percentage of agreement between all possible pairs of raters. To do this, one divides the number of agreements by the total number of ratings/observations on a scale and/or subscale (Baer, 1977). Although there is some debate regarding the acceptability of percent agreement as an estimate of inter-rater reliability, some consensus has been reached that an average of 70% is necessary, 80% is a considered adequate, and 90% is considered good (House, House, & Campbell, 1977).

While calculating the percentage of agreement across raters is simple and provides an estimate of a measure's inter-rater reliability, it does not take into account the amount of agreements expected based upon chance alone (Cohen, 1960). Thus, one of the most common methods for assessing inter-rater reliability is using Cohen's kappa (k), which is a statistic that does correct for chance (Cook & Beckman, 2006; DeVellis, 2005). The other frequently used measure to assess inter-rater reliability is Intra-Class Correlations (ICCs; Shrout & Fleiss, 1979).

Validity

Validity refers to the extent to which a scale measures what it intends to measure (McGartland-Rubio, 2005; McGartland-Rubio et al., 2003; Nunnaly & Bernstein, 1994). Similar to reliability, many types of validity exist, such as construct validity, predictive validity, translation validity, concurrent validity, and so on (Netemeyer et al., 2003; Nunnaly & Bernstein, 1994). However, construct validity is one of the most critical types of validity for a scale developer to assess (Netemeyer et al., 2003). McGartland-Rubio (2005) defines construct validity as, "The extent to which an item or measure accurately represents the proposed construct" (p. 495). Construct validity cannot be directly assessed; rather, researchers must gather evidence for other types of validity, such as content validity (AERA, 1999; Crocker & Algina, 1986; Haynes, Richard, & Kubany, 1995). While factor analysis is the most highly and commonly recommended procedure for assessing the validity of a scale, it is strongly recommended that the scale be administered to a sample size of at least 300 (Comrey, 1973; DeVellis, 2012; Worthington & Whittaker, 2006). There are other methods for assessing a scale's validity for sample sizes of less than 300, such as recruiting a panel of experts to assess content validity.

Content validity. The content validity of a measure is important because it tells us how well the items on a measure represent a construct's theoretical domain (Messick, 1993; Nunnally & Bernstein, 1994). One of the ways a scale's content validity can be examined is by enlisting a panel of experts to review the initial item pool, which is also a distinct step recommended in the scale development process (DeVellis, 2012; Haynes et al., 1995; McGartland-Rubio, 2005; McGartland-Rubio et al., 2003). Two distinct populations should be considered for inclusion in the expert panel: (1) "lay experts" for the population the scale is being developed for and (2)

experts who have published in the field (McGartland-Rubio, 2005). Thus, for the purposes of the current study, both school personnel serving on or consulting with a PST and experts in the fields of problem-solving process/teams, ABA, and FBA, would be recruited for the expert panel. A minimum of three experts per group (*n*=6) is recommended, with a maximum of 10 experts per group (*n*=20; McGartland-Rubio, 2005). Several previous educational/psychological scale development and validation studies have utilized this method, with the number of expert reviewers recruited ranging between 5, 6, and 24 expert reviewers, respectively (Coplan & Rubin, 1998; Eliasson et al., 2006; Feeney-Kettler et al., 2011).

In order to assess the content validity of a measure, the panel of experts must rate each item based on three elements: (1) which of the defined factors, if any, the item belongs to, (2) the item's representativeness of the assigned factor (response options 1-4; 1=not representative and 4=representative), and (3) the clarity of the item (1=item is not clear and 4=item is clear; McGartland-Rubio, 2005). See Table 1 below for an example.

Figure 1. Content Validity Assessment Form

#	Item	Factor					Representativeness				Clarity			
1		1	2	3	4	5	1	2	3	4	1	2	3	4
2		1	2	3	4	5	1	2	3	4	1	2	3	4

Adapted from the National Association of Social Workers, Inc. (as cited in McGartland-Rubio, 2005)

Once this information is collected from participating reviewers, responses may be analyzed to assess the factorial validity index (FVI) and content validity index (CVI) of the items and measure. The FVI of a measures the percentage of expert reviewers who correctly assign the items to their intended factors, whereas the CVI measures how well the reviewers believe the items represent the factor or construct. The FVI and CVI should both reach 80% in order to indicate strong content validity (McGartland-Rubio, 2005; McGartland-Rubio et al., 2003).

Summary

The need for implementation of EBIs to address behavioral disorders among children in the school setting continues to increase (Costello et al., 2004; Hoagwood et al., 2001). Furthermore, many organizations, researchers, advocates, and even legislation have called for an emphasis on the uptake of EBIs in school settings (IDEA, 2004; Kratochwill & Stoiber, 2000; NCLB, 2002; APA Task Force, 2008). Despite efforts to make adoption and implementation of EBIs more efficient (e.g., WWC), the rate with which schools, including school-based PSTs, do adopt and implement EBIs is failing to match the need with little evidence as to why (Walker, 2004). Similar to other fields, scholars and practitioners in educational fields need to examine systems and structural factors, specifically factors related to the functional behavior of PSTs, related to this research-to-practice gap (Burns et al., 2008b; Glisson, 2007). The purpose of this study was be to modify the PDC for application with school-based PSTs, and conduct an initial exploratory investigation of the revised PDC's psychometric properties. Specifically, the items on the PDC-S will be revised and/or added to; content validity will be examined using the judgment and feedback of a panel of expert reviewers; inter-rater reliability will be tested; and the PDC-S's test-retest reliability will be investigated.

Research Questions and Predictions

Below is a description of the research questions and hypotheses. The hypotheses are guided by a collaborative consideration given to best practices in scale development, recommended guidelines for establishing scale reliability and validity, and procedures followed in similar development and validation studies. This study will be guided by the following research questions.

1. How strong is the content validity of the *Performance Diagnostic Checklist-Schools (PDC-S)?*

Hypothesis 1: The PDC-S will exhibit strong content validity with the overall percentage of items correctly classified by expert reviewers (the factorial validity index; FVI) reaching ≥80%, and the representativeness of the correctly classified items (content validity index; CVI) reaching ≥80%.

It is recommend in the literature that both FVIs and CVIs reach at least 80% to be considered adequate (McGartland-Rubio, 2005; McGartland-Rubio et al., 2003).

2. Is the inter-rater reliability acceptable for the PDC-S relative to what individual PST members identify as potential barriers to the adoption and implementation of EBIs?

Hypothesis 2: Inter-rater reliability will vary somewhat within PSTs as each member has his or her own perception of the team's processes. However, it is hypothesized that the average percent agreement among raters within each PST will be ≥ 70%.

In a similar study, Iwata, DeLeon, and Roscoe (2013) assessed the inter-rater reliability (agreement) of a tool similar to the PDC-S (the *Functional Analysis Screening Tool* [FAST]), and found an overall percent agreement of 71.50%. Percent agreement values ranging from 34.67% to 78.56% were also found for similar measures in a study conducted by Barton-Arwood and colleagues (2003).

3. Are within-group percent agreements higher than between-group percent agreements on the PDC-S in regards to inter-rater reliability?

Hypothesis 3: Percent agreements will be higher within-groups than between-groups for inter-rater reliability.

4. What is the test-retest reliability of the PDC-S?

Hypothesis 4: The test-retest reliability of the PDC-S will be high, with correlation coefficients of 0.80 or higher.

It is recommended that all reliability coefficients be at least 0.80 (Webb et al., 2006). Test-retest reliability of the QABF was calculated for subscales and the test as a whole using the Pearson product-moment coefficient, and yielded coefficients ranging from 0.795 to 0.990 (Paclawaskyj et al., 2000).

CHAPTER 3

Method

This study was conducted in three phases: (1) the PDC-S was adapted from its original version into a tool applicable for use with school-based PSTs, (2) two panels of experts reviewed the PDC-S to determine the content validity of the adapted PDC-S, and (3) the adapted PDC-S was administered to team members within a sample of PSTs two times to determine both inter-rater reliability and test-retest reliability. Thus, this section will be discussed within the context of these three research phases.

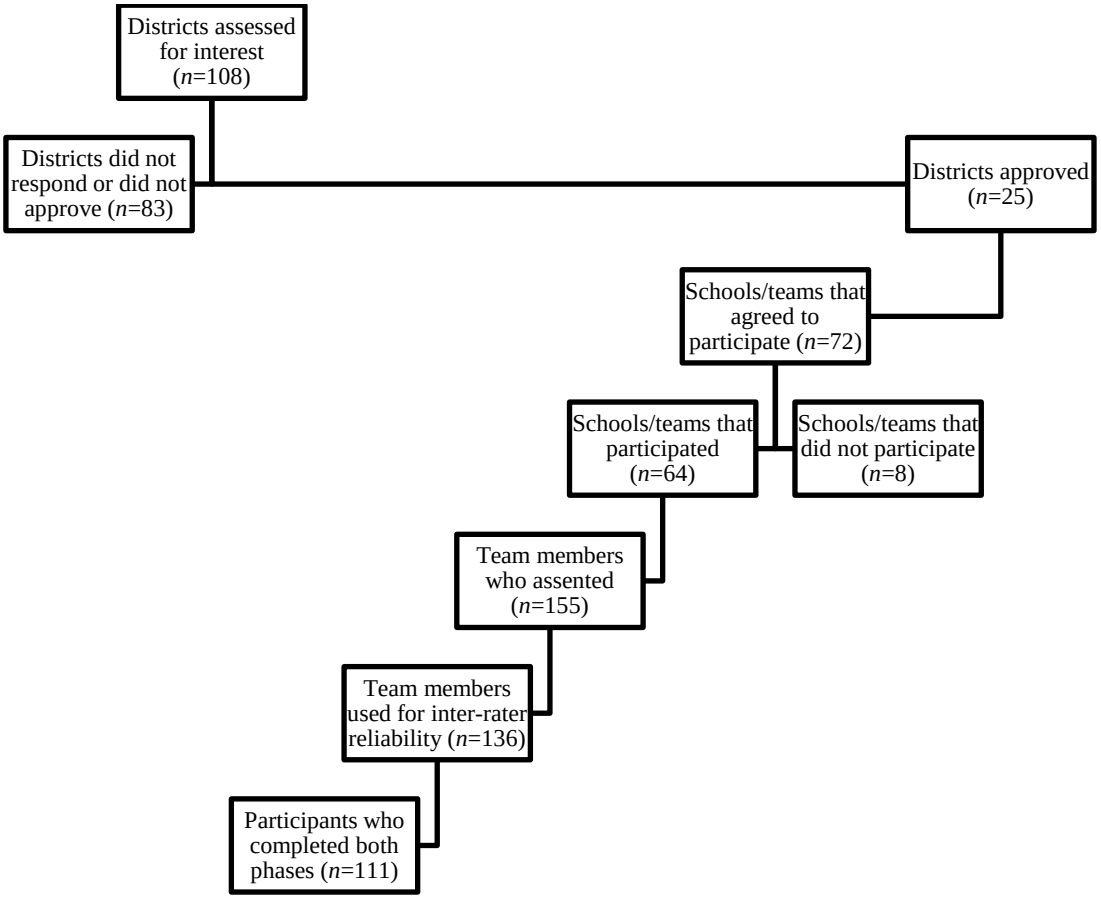
Participant Recruitment

Phase 2: Expert panel review. A list of experts to recruit was generated in collaboration with the project director and principal investigators from the LIST Project to include members of journal editorial boards focused in the area of ABA/FBA and schools (i.e., *School Psychology Revi*ew and *Journal of Applied Behavior Analysis* [JABA]). Only professionals with training in and understanding of ABA and FBA principles were asked to participate. A recruitment email was sent to 86 eligible experts. Seventeen experts elected to participate, exceeding the target sample size of 15 expert reviewers. Thus, the final n =17 exceeds the minimum recommendation of n = 6 (McGartland-Rubio, 2005; Netemeyer et al., 2003). For phase 2b, all members (n = 9) of the LIST Project Advisory Committee were recruited to participate, with four electing to participate.

Phase 3: Administration of the PDC-S. The sample for the third phase of the current research study was determined using two selection criteria. The first requirement for inclusion in the study was that participants must be members of a PST that addresses behavioral concerns for students grades 4K through fifth (elementary-level). The LIST Project focuses only on teams

within elementary-level schools, thus the PDC-S was validated with the population of its intended use. Another selection criterion was that only teams and team members not currently participating in the LIST Project be recruited for the current study. This criterion was established to protect the integrity of the LIST Project data and findings. Figure 2 below summarizes the sampling procedure and percentage of participating schools and team members.

Figure 2. Participating Schools Flow Chart



A total of 108 school districts across two Midwestern states were contacted and assessed for interest in participation. About a quarter of the districts contacted approved the current study and allowed recruitment of PST members from their schools (*n*=25). A total of 72 schools agreed to participate, however only 64 of those schools had team members interested in

participating. A total of 155 team members participated in this phase of the study. Because it was not required that the entire PST either agree or decline to participate, there were some cases where only one team member from a PST participated. Thus, 136 participants' responses were used to assess the inter-rater reliability of the PDC-S. Several participants failed to complete the PDC-S during the second administration, leaving 111 response sets to be used for assessing the test-retest reliability of the PDC-S.

Participants

Phase 1: Adaptation of the PDC. In addition to the primary researcher, the co-Principal Investigators and Director of the LIST Project worked collaboratively to revise the original PDC into the PDC-S. Members from the LIST Project Advisory Committee also assisted during this phase of the study.

Phase 2a: Expert panel review. A total of 17 professionals participated in this phase of the study as expert reviewers, exceeding the recommended minimum number of expert reviewers (n = 6) based on best practices in the scale development literature (DeVellis, 2012; McGartland-Rubio, 2005; Netemeyer et al., 2003). The panel of expert reviewers recruited for phase two of this study included a variety of professionals in fields related to psychology and education, including researchers (47.1%), practitioners (29.4%), and those who identified themselves as both researchers and practitioners (23.5%). Experts represented a diverse range of ages, experiences, and education (see Table 2). The gender of expert reviewers was almost equal, with females comprising 52.9% of the participants and males making up 47.1%. The highest percentage of expert reviewers fell within the ages of 35 to 44 (41.1%), followed by the 55 and older age range (23.5%), followed by the less than 35 and the 45 to 54 age range (17.7% respectively). The majority of expert reviewers were Caucasian (88.2%), followed by African

American or Black; Asian American; and Hispanic or Latino/a (5.9% respectively). The highest percentage of expert reviewers had 20 or more years of experience working in the education and/or mental health field(s) (47.1%), followed by 6 to 10 years and 16 to 20 (23.5% respectively), then followed by 11 to 15 years (5.9%). All expert reviewers had at least some training in ABA/FBA principles and procedures; the majority had five or more years (52.9%), followed by three to four years (23.5%), then followed by one to two years and less than one year (11.8% respectively). The highest percentage of expert reviewers had obtained 20 or more years in utilizing principles and procedures of ABA/FBA (35.3%), followed by 6 to 10 years (29.4%), then followed by one to five years and 11 to 15 years (17.7% respectively). In regards to the highest academic degree obtained, the majority of expert reviewers possessed a doctorate (76.5%) with the remaining expert reviewers possessing a Master's or Specialist degree (23.5%).

Table 1

Phase 2a Expert Reviewer Demographics (n = 17)

Variable	*n*	%
Profession		
Researcher	8	47.1
Practitioner	5	29.4
Both	4	23.5
Gender		
Male	8	47.1
Female	9	52.9
Age		
Less than 35	3	17.7
35-44	7	41.1
45-54	3	17.7
55+	4	23.5
Ethnicity[1]		
African American or Black	1	5.9
American Indian or Alaskan Native	0	0.0
Asian American	1	5.9
Caucasian or White	15	88.2
Hispanic or Latino/a	1	5.9

Native Hawaiian or other Pacific Islander	0	0.0
Other	0	0.0
Number of years working in education/mental health field(s)	0	0.0
0-5	4	23.5
6-10	1	5.9
11-15	4	23.5
16-20	8	47.1
20+		
Number of years training in ABA/FBA		
None	0	0
Less than 1 year	2	11.8
1-2 years	2	11.8
3-4 years	4	23.5
5+ years	9	52.9
Number of years experiencing utilizing ABA/FBA	0	0.0
None	0	0.0
Less than 1 year	3	17.7
1-5 years	5	29.4
6-10 years	3	17.7
11-15 years	0	0.0
16-20 years	6	35.3
20+		
Highest Academic Degree		
High School Diploma	0	0.0
Bachelor's	0	0.0
Master's/Specialist	4	23.5
Doctorate	13	76.5
Other	0	0.0

[1] Participants were instructed to select "all that apply"

Phase 2b: Second expert panel review. Members from the LIST Project Advisory Committee were recruited to conduct a second expert panel review. Participants for this phase included a total of four professionals. The majority of participants identified themselves as researchers (75.0%) while one reviewer identified themselves as both a practitioner and researcher (25.0%). At least half of the participants were male (50.0%) and one female (25.0%). One participant did not respond to this question. Of the three expert reviewers who identified their age, two were 55 or older (50.0%) and one was between the ages of 45 and 54 (25.0%).

Regarding ethnicity, all participants who responded selected "Caucasian" (75.0%). The majority of participants have been working in the field(s) of education and/or mental health for 20 or more years (75.0%), followed by 16 to 20 years (25.0%). Participants' training in ABA and FBA varied with equal numbers of experts indicating the following number of years of training received: none (25.0%), one to two years (25.0%), three to four years (25.0%), and five or more years (25.0%). Three participants identified the number of years of experience they have in utilizing procedures of FBA and ABA with half selecting 20 or more years (50.0%) and one selecting 16 to 20 years (25.0%). All three of the four expert reviewers who responded to the final item indicated that the highest academic degree they had obtained was a doctorate (75.0%).

Table 2

Phase 2b Expert Reviewer Demographics (n =4)

Variable	*n*	%
Profession		
Researcher	3	75.0
Practitioner	0	0.0
Both	1	25.0
Gender[1]		
Male	2	50.0
Female	1	25.0
Age[1]		
Less than 35	0	0.0
35-44	0	0.0
45-54	1	25.0
55+	2	50.0
Ethnicity[1]		
African American or Black	0	0.0
American Indian or Alaskan Native	0	0.0
Asian American	0	0.0
Caucasian or White	3	75.0
Hispanic or Latino/a	0	0.0
Native Hawaiian or other Pacific Islander	0	0.0
Other	0	0.0
Number of years working in education/mental health field(s)		

0-5	0	0.0
6-10	0	0.0
11-15	0	0.0
16-20	1	25.0
20+	3	75.0
Number of years training in ABA/FBA		
None	1	25.0
Less than 1 year	0	0.0
1-2 years	1	25.0
3-4 years	1	25.0
5+ years	1	25.0
Number of years experience utilizing ABA/FBA[1]	0	0.0
None	0	0.0
Less than 1 year	0	0.0
1-5 years	0	0.0
6-10 years	0	0.0
11-15 years	1	25.0
16-20 years	2	50.0
20+		
Highest Academic Degree[1]		
High School Diploma	0	0.0
Bachelor's	0	0.0
Master's/Specialist	0	0.0
Doctorate	3	75.0
Other	0	0.0

[1] $n = 3$ for this item

Phase 3: Administration of the PDC. Participants were recruited from elementary-level schools in two Midwestern states. A total of 25 school districts approved recruitment of participants from their schools, and some districts had participants from more than one school/PST. The distribution of participants across districts is summarized in Table 3 below. Participants were various educational professionals who served as members on their schools' problem-solving teams. Participants held a variety of educational positions typical for PST composition. The most represented position was that of general classroom teachers (27.1%), followed by school guidance counselors (15.5%), then special education teachers (14.8%), principals (12.3%), and school psychologists (8.4%). There was minimal but equal

representation across the following positions: reading specialist, behavior analyst/specialist, speech and language pathologist, and school social worker (4.0% respectively). A few paraprofessionals also participated (3.0%). There were some other roles represented to a minimal degree, including assistant principal, English Language Learner (ELL) classroom teacher, and student services director; these participants fell under the "other" category and comprised 9.7% of the sample.

As one might expect of the Midwest, participating PST members were largely female (89.7%) and white (96.8%;National Center for Education Statistics, 2012) . The age of participants was well-distributed across ranges; the highest percentage of participants were in the 45-54 age range (31.6%), closely followed by "less than 35" and 35-44 age ranges (29.7% respectively). A handful of participants fell in the 55+ age range (9.0%). The majority of participants held Master's or Specialists degrees (74.2%), followed by Bachelor's degrees (21.9%), and then Doctorate degrees (1.9%).

The highest percentage of participants have been working in the field of education for 0-5 years (21.1%), followed by 11-15 years (18.4%), and then 6-10 years (16.5%). An equal number of participants have been working in the field of education for 16-20 years and 21-25 years (15.1%), followed by 26-30 years (7.2%), and finally 31+ years (6.6%). The majority of participants had only been working at their current school for 0-5 years (52.3%), followed by 6-10 years (20.6%), then 11-15 years (11.6%), 20+ years (9.0%), and 16-20 years (6.5%). The majority of participants have served on their schools' PSTs for 3+ years (62.0%), followed by 1-2 years (20.6%), then 2-3 years (16.8%), 6-12 months (22.0%), 3-6 months (7.0%), and finally less than three months (2.6%).

Only one participant held a certification in behavior analysis (i.e., is a Board Certified Behavior Analyst [BCBA]). However, participants reported having some training in and experience with principles of Applied Behavior Analysis (ABA) and methods of Functional Behavioral Assessments (FBA). The majority of participants had less than one year of formal training in ABA and/or FBAs (66.5%), followed by 1-2 years (18.1%), then 3-4 years (8.4%), and finally 5+ years (5.2%). When asked how many years of experience participants had in utilizing principles of ABA and methods of FBA, the majority responded with "less than one year" (56.1%), followed by 1-3 years (20.6%), then 4-7 years (7.7%), and 8-12 years and 13+ years (6.5% respectively).

Table 3

PST Member Demographic Information (n = 155)

Variable	*n*	%
School District		
District A	2	1.3
District B	3	1.9
District C	23	14.8
District D	4	2.6
District E	26	16.8
District F	3	1.9
District G	2	1.3
District H	12	7.7
District I	2	1.3
District J	1	0.6
District K	3	1.9
District L	1	0.6
District M	9	5.8
District N	5	3.2
District O	2	1.3
District P	16	10.3
District Q	8	5.2
District R	12	7.7
District S	2	1.3
District T	11	7.1
District U	2	1.3
	2	1.3

District V	1	0.6
District W	1	0.6
District X	2	1.3
District Y		
Main role in school		
General education classroom teacher	42	27.1
School Counselor	24	15.5
Special education classroom teacher	23	14.8
Principal	19	12.3
School Psychologist	13	8.4
Reading Specialist	4	2.6
Speech and Language Pathologist	4	2.6
Behavior Analyst/Specialist	4	2.6
School Social Worker	4	2.6
Paraprofessional	3	1.9
Other	15	9.7
Gender		
Male	16	10.3
Female	139	89.7
Age		
Less than 35	46	29.7
35-44	46	29.7
45-54	49	31.6
55+	14	9.0
Ethnicity		
African American or Black	0	0.0
American Indian or Alaskan Native	0	0.0
Asian American	2	1.3
Caucasian or White	150	96.8
Hispanic or Latino/a	1	0.6
Native Hawaiian or other Pacific Islander	0	0.0
Other	2	1.3
Grade(s) currently assigned to/serving[1]		
Pre-Kindergarten (4K)	41	26.5
Kindergarten	100	64.5
First grade	112	72.3
Second grade	97	62.6
Third grade	102	65.8
Fourth grade	99	63.9
Fifth grade	86	55.5
Number of years teaching or working in education[4]	32	21.1
0-5	25	16.5
6-10	28	18.4
11-15	23	15.1

16-20	23	15.1
21-25	11	7.2
26-30	10	6.6
31+		
Number of years teaching or working at current school	81	52.3
0-5	32	20.6
6-10	18	11.6
11-15	10	6.5
16-20	14	9.0
21+		
Number of years serving on problem-solving team[3]	4	2.6
Less than 3 months	7	4.5
3-6 months	22	14.2
6-12 months	32	20.6
1-2 years	26	16.8
2-3 years	62	40.0
3+ years		
Number of years training in ABA/FBA[4]		
Less than 1 year	103	66.5
1-2 years	28	18.1
3-4 years	13	8.4
5+ years	8	5.2
Number of years experiencing utilizing ABA/FBA[5]	87	56.1
Less than 1 year	32	20.6
1-3 years	12	7.7
4-7 years	10	6.5
8-12 years	10	6.5
13+		
Type of teaching/other license held[1]		
Regular teaching license	116	74.8
Administrator certificate	25	16.1
School counselor certificate	15	9.7
School psychology license	14	9.0
Emergency license	3	1.9
School social worker license	3	1.9
Speech/language pathologist license	2	1.3
Other	9	5.8
Highest Academic Degree[2]		
Bachelor's	34	21.9
Master's/Specialist	115	74.2
Doctorate	3	1.9
Other	2	1.3

[1] Participants were instructed to select "all that apply"

[2] *n* = 154 for this item
[3] *n* = 153 for this item
[4] *n* = 152 for this item
[5] *n* = 151 for this item

Participating Problem-Solving Teams

In addition to demographic information collected regarding the participating team members, information regarding the teams as a whole was collected from each participating team member. Teams were asked to identify their team composition; a large majority of teams listed their principal as a permanent team member (90.3%). The majority of team members also listed the following as permanent team members: school guidance counselor (81.3%), general classroom teacher (76.1%), and special education teacher (74.8%). Nearly half of team members noted a school psychologist as a permanent team member (49.7%). Other identified team members include school social workers (25.8%), reading specialists (23.2%), speech and language pathologists (20.0%), paraprofessionals (16.1%), behavior analyst/specialists (11.0%), and assistant principal (6.5%). Some participants indicated "other" team members not previously listed (23.9%); examples of these include Title I teachers, deaf educators, and specials teachers (i.e., art teachers, physical education teachers, etc.). The highest percentage of teams had a total of five permanent team members (21.1%), followed by seven (18.4%), and then six (16.5%). Some teams had eight permanent team members (11.8%), four team members (8.6%), three team members (7.2%), two permanent team members (2.6%), and nine (1.3%). Several teams had 10 or more permanent team members (11.2%), and a couple of teams selected "other" indicated that the number of team members vary from week to week (1.3%).

Most teams met biweekly (i.e., every two weeks or twice per month; 43.9%), followed by teams that met weekly (29.0%), then monthly (i.e., once per month; 16.8%), once per quarter

(3.8%), and finally once per semester (0.6%). Some teams indicated that they met on an "as needed" basis; these responses fell under the "other" category, comprising 5.1% of participant responses. Almost half of the teams had been meeting for 5+ academic years (46.5%). Other teams had been meeting for 3-4 academic years (20.0%), followed by 1-2 academic years (17.4%), then 6-9 months (10.3%), 0-3 months (3.2%), and 3-6 months (1.9%). All but one of the participating schools served Kindergarten (99.4%). The next most commonly served grade was first (96.8%), followed by second (95.5%), then third (94.2%), fourth (92.3%), and finally fifth (81.9%).

Table 4

Problem-Solving Team Information

Variable	*n*	%
Team composition[1]		
Principal	140	90.3
School guidance counselor	126	81.3
General education classroom teacher	118	76.1
Special education classroom teacher	116	74.8
School Psychologist	77	49.7
School Social Worker	40	25.8
Reading Specialist	36	23.2
Speech and Language Pathologist	31	20.0
Paraprofessional	25	16.1
Behavior Analyst/Specialist	17	11.0
Assistant Principal	10	6.5
Other	37	23.9
Longevity of team[2]		
0-3 months	5	3.2
3-6 months	3	1.9
6-9 months	16	10.3
1-2 academic years	27	17.4
3-4 academic years	31	20.0
5+ academic years	72	46.5
Frequency of team meetings[2]		
Weekly	45	29.0
Biweekly	68	43.9
Monthly	25	16.8

Once per quarter	5	3.8
Once per semester	1	0.6
Other (e.g., "as needed")	8	5.1
Number of permanent team members[3]		
2	4	2.6
3	11	7.2
4	13	8.6
5	32	21.1
6	25	16.5
7	28	18.4
8	18	11.8
9	2	1.3
10+	17	11.2
Other (e.g., "It varies from week to week")	2	1.3
Grade(s) served[1]		
Kindergarten	154	99.4
First grade	150	96.8
Second grade	148	95.5
Third grade	146	94.2
Fourth grade	143	92.3
Fifth grade	127	81.9

[1] Participants were instructed to select "all that apply"
[2] *n* = 154 for this item
[3] *n* = 152 for this item

Measures

A total of five measures were administered during the current study. All measures were administered and completed by participants online via University of Wisconsin-Madison's survey host, *Qualtrics*.

Performance Diagnostic Checklist-Schools. The original form of the PDC (see Appendix A), which was created for application in the field of OBM, was adapted during the initial phase of this study into the PDC-S (this process and the measure itself will be discussed in greater detail below). The adapted PDC-S was then sent to two panels of experts for review during phase two. After revisions were made to the final PDC-S (see Appendix C) based on the

expert reviewers' feedback, it was administered to recruited PST members twice during phase three of the current study.

Phase 2: Expert panel review. Several measures were utilized during the second phase of the current study including a demographic questionnaire and the Expert Review Form, in addition to the PDC-S.

Expert review demographic questionnaire. This questionnaire was designed to gather information about the panel of expert reviewers recruited to participate in phase two. The demographic information gathered included the gender, age, ethnicity, occupation, number of years in occupation, and highest degree attained by all recruited experts (see Appendix D). Because only professionals with training, experience, and knowledge of ABA and FBA principles were recruited for participation, this form also asked participants to indicate (a) the years of training they received in ABA/FBA and (b) the years of experience (e.g., teaching or practicing) they have in utilizing ABA/FBA principles and procedures.

Expert review form. Each expert reviewer who elected to participate in the current study was sent an email including a link to the Expert Review Form (see Appendix E). This form asked expert reviewers to indicate (1) which of the defined factors, if any, the item belongs to, (2) the item's representativeness of the assigned factor (response options 1-4; 1=not representative and 4=representative), and (3) the clarity of the item (1=item is not clear and 4=item is clear; McGartland-Rubio, 2005; McGartland-Rubio et al., 2003). In addition, there was a comment box at the end of each item for expert reviewers to make any comments regarding the content and/or clarity of the item.

Phase 3: Administration of the PDC. During the third phase of the current study the following measures were administered to participants, in addition to the PDC-S: the PST information sheet and the PST member demographic questionnaire.

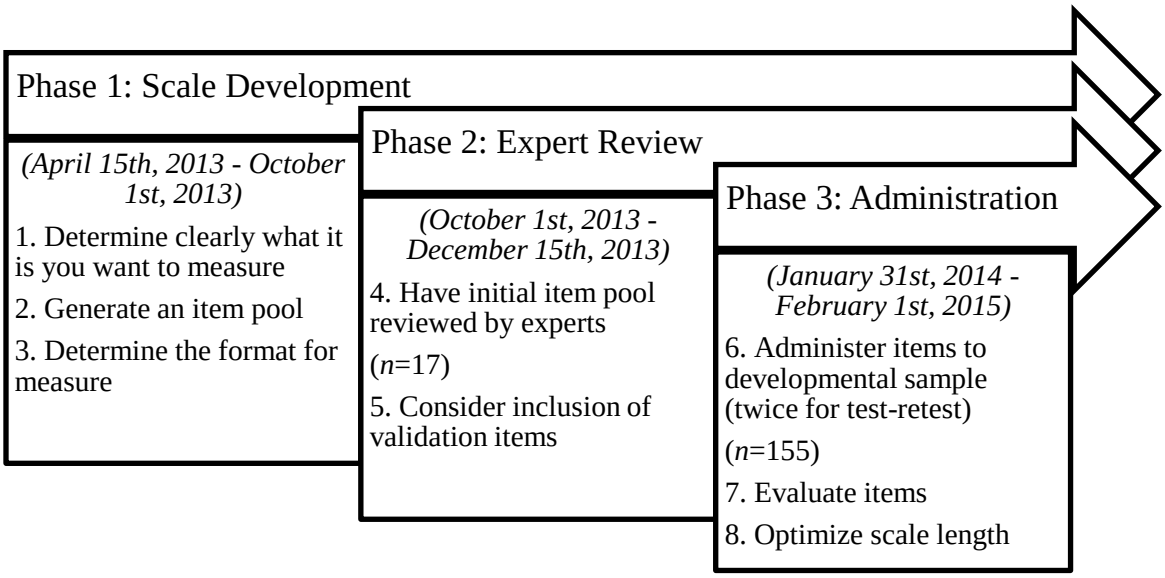
Problem-solving team information sheet. Each participant completed an information sheet regarding their team's characteristics (see Appendix F). The information gathered from each PST member included the number of team members, team composition, grades served, and number of months/years the team had been operating.

PST member demographic questionnaire. In addition to gathering information about each PST, each team member was also asked to complete a demographic questionnaire (see Appendix G). This questionnaire asked PST members their age, gender, ethnicity, number of years practicing, current role on the PST, grades served, number of years at the current school, number of years serving on the PST, years of training in FBA/ABA, years utilizing principles/procedures of FBA/ABA, type of license, and highest level of education attained.

Procedures

As described above, this project was a multi-phase process beginning with the adaptation of the PDC from its original form for use with PSTs. After the PDC was adapted, a panel of expert reviewers was enlisted to determine the appropriateness of the items. Once the expert reviewers' feedback was incorporated into the PDC-S, it was administered to recruited PST members twice to test its reliability – both test-retest and inter-rater. Both administrations took place between the months of January 2014 and February 2015. Please see Figure 3 below for an overview of the timeline and procedures.

Figure 3. Timeline of Study



Phase one: Adaptation of the PDC. As discussed above, the process of developing the PDC-S was guided by DeVellis' (2012) proposed scale development process: (1) determine clearly what you want to measure, (2) generate an item pool, (3) determine the format for measure, (4) have initial item pool reviewed by experts, (5) consider inclusion of validation items, (6) administer items to a development sample, (7) evaluate the items, and (8) optimize scale lengths. Although the PDC was adapted from its OBM form for application with school PSTs, rather than developed originally, the researcher completed the steps outlined above to ensure the development of a valid and reliable tool.

Step 1: Determining what is to be measured. The decision regarding what was to be measured was fulfilled by the LIST Project within which this study resides. The goal of the PDC-S, as asserted by Kratochwill and Asmus (2011), is to identify the impediments to consistent adoption and implementation of EBIs (p. 16). Furthermore, by assigning items to one of the four broad areas, the goal is more specific in that impediments will be identified by

training/knowledge; task clarification and prompting; resources, materials & processes; and performance feedback/consequences & outcomes. Ultimately, the aim is for the identification of impediments within specific areas to generate functional hypotheses regarding PSTs' use of EBIs that may be tested during subsequent phases of the LIST Project.

Step 2: Generating an item pool. Items on the PDC-S were developed based on several primary sources: theoretical principles of ABA; procedural guidelines of FBA; the original PDC (Austin, 2000); the PDC-HS (Carr et al., 2013); best practices in PSTs and processes; alignment with the DORA-II; and literature and practical knowledge of EBI selection and implementation. In addition to consulting these sources, the primary researcher worked in collaboration with the principal investigators from the LIST project whom have extensive knowledge of ABA and FBA to generate the initial item pool.

This process started by taking the original PDC and changing the wording of items to better fit the context of school-based PSTs. While this was occurring, additional items were included that covered more specific characteristics typically present within a PST, according to the literature. The four areas within the PDC-S were also slightly revised in order to better fit the context of PSTs. These four areas were intended to capture four general behavioral functions of PSTs. In other words, if a team produced an overall low score on the PDC-S, their area scores would help indicate why PSTs were not functioning optimally. For example, if a PST had a low score in the "Training/Knowledge" area but high scores in the other three areas, that might suggest that the reason the PST wasn't functioning well is due to a lack of training in effective PST processes and not due to other reasons (e.g., lack of task clarification and prompting [area 2]).

A thorough review of the DORA-II (another assessment tool utilized by the LIST Project) was also conducted by the research team so that items on the PDC-S and the DORA-II aligned with one another. This was an important step since the DORA-II focuses on evidenced-based problem solving steps. Data from the PDC-S provides information regarding what each team endorses as happening in the team process, whereas data from the DORA-II provides information about what is actually happening based on direct observation by trained coders. The process of generating an initial item pool for the PDC-S took several months of ongoing collaboration between the LIST Project co-PIs, the LIST Project Manager, and the primary researcher for the current study.

Step 3: Determining the format for measure. The format for the structure of how the PDC-S is presented was determined by both the previous purpose and version of the PDC, as well as the researchers on the LIST Project. For each item included in the PDC-S, PST members were asked to answer three questions: (1) whether the item was currently in place within their PST (possible answers: *yes* or *no*), (2) if participants answered "no" to the previous question, they were asked to select one of five possible reasons for why they believe that particular item is not in currently in place on their team (possible answers: *other needs take priority over this activity/action during team meetings, limited knowledge/experience with this activity, lack of communication/direction provided to the team regarding this activity/action, limited support/acknowledgement from administration/team for this action,* or *unable to devote time to support this activity/action*), and (3) how important they viewed the item (possible answers: *very, somewhat,* or *not really*). It was important to include these questions in the PDC-S because intra-team agreement on responses can yield crucial information for the development of testable functional hypotheses, as well as function-based interventions.

Step 4: Have initial items reviewed by experts. After the initial version of the PDC-S was developed, it was administered to two panels of expert reviewers during the second phase of the present study. As discussed above, the purposes for expert review were to (a) rate the relevance of each item to the factor, (b) evaluate the clarity of each item, and (c) indicate any missed items that could tap into the construct being measured (DeVellis, 2012). Thus, reviewers were provided with a working definition of the construct, asked to assign each item to its appropriate category based on these definitions, make revisions to the wording of items, and suggest any additional items to be included.

After the initial pool of items on the PDC-S was reviewed by a panel of experts, the items were revised based on the results and feedback from those experts. The primary researcher gathered the results and comments from the reviewers and made suggested edits to the items and PDC-S as a whole. These recommended edits were made based on the FVI and CVI ratings of each item, as well as the clarity ratings. For example, if the FVI of an item was less than 50%, and there was another area for which the majority of expert reviewers assigned an item to, it was recommended that that item be moved to that respective area. Additionally, if an item had mixed results (i.e., no clear majority assignment to one specific area), revisions to the wording of that item were made in order to make it more reflective of the area within which it was originally intended to fall. Revision to the wording of items was largely based off of the comments provided from expert reviewers.

The LIST Project co-PIs and Project Manager then reviewed the recommended changes and determined whether or not to incorporate them. These individuals also made their own recommendations for edits based on the results of Phase 2a. After several rounds of reviews and edits, the final pool of items was established. In order to determine whether the revisions were

sufficient in improving the validity and clarity of the items, a second expert panel review was conducted (i.e., Phase 2b). The same review and revision process of the PDC-S items took place following the completion of this second expert panel review, thus creating the final version of the PDC-S. For some items a strong FVI was not present for any area, indicating that those items were not fully representative of any one area. However, the research team believed the information to be gathered from those items to be important for the overall intent of the PDC-S, so they were assigned to the area which most expert reviewers assigned them to.

Step 5: Consider inclusion of validation items. The two major types of validity items that are recommended for inclusion in a scale examine the social desirability of informants' answers and the construct validity of a measure (DeVellis, 2012). Due to the nature of the PDC-S, it is not appropriate to include a social desirability scale. However, the following techniques were utilized to reduce the likelihood of socially desirable responses: self-administration of the PDC-S and anonymity (Netemeyer et al., 2003). Enlisting a panel of expert reviewers to assign each item to the most appropriate construct assessed the content validity of the PDC-S; this took place during the second phase of the current study, which is described in greater detail below.

Step 6: Administer items to development sample. The administration of the PDC-S items to a development sample took place during phase three of the current study, described in greater detail below.

Step 7: Evaluate items. Once the PDC-S was administered to the recruited sample of PSTs, generated scores were used to evaluate the items. During this step, median scores for each item and category were calculated using the results from the administration of the PDC-S in Phase three. In addition, ranges and confidence intervals were identified. This step in adapting the PDC-S allowed researchers to finalize methods of scoring and interpreting PDC-S results.

Step 8: Optimize scale lengths. Based on the results of the former step of the development phase, as well as the results from the expert panel review during phase two, certain items were removed from the PDC-S to create the optimal length and include only the most valid, clear, and reliable items.

Phase 2: Expert panel review. To test the content validity of the adapted PDC-S, the researcher sought the expertise of two panels of experts in the fields of education and/or mental health with experience in ABA and FBA principals and procedures. The researcher obtained a waiver of signed consent from the IRB (see Appendix H for the most recent IRB approval), allowing all steps of the process to be completed online via methods of e-mail and *Qualtrics*. Identified experts were sent an email with a link to an overview of the study parameters (see Appendix I) up to two times. At the end of this document was the statement, "I have read, understood, and printed a copy of, the above consent form and desire of my own free will to participate in this study," to which participants could select "no" and decline to participate or "yes" in which case they were directed to the instructions for completing the survey. The survey included definitions of the four PDC-S factors and a list of all the PDC-S items developed by the researchers. The panel was asked to indicate which PDC-S category each item best represented. Panel members were also asked to suggest any additional items they believed would be essential to measuring the construct. The primary researcher, along with the LIST co-principal investigators and project manager, reviewed the results from the expert panel to determine appropriate revisions to be made to the PDC-S. After these revisions were incorporated into the PDC-S, a second panel of expert reviewers was recruited to repeat the process outlined above. Results from this second expert panel review again informed revisions made collaboratively between the primary researcher and LIST Project personnel to create a final version of the PDC-

S to be administered to a developmental sample in Phase 3. Once a final version of the PDC-S was established, the researcher created an online version of the PDC-S using *Qualtrics.* All participating expert reviewers were entered in a drawing to win a $50 gift card. A total of 4 expert reviewers were randomly selected as the winners of the gift cards.

Phase 3: Administration of the PDC-S. After the expert panel reviewed the PDC-S and the researcher incorporated expert reviewers' feedback into the adapted PDC-S, it was administered to recruited PSTs. It was not required that all members of a single PST complete the PDC-S. Thus, each individual member from district- and school-approved PSTs was sent an email that included a link to an overview of the study parameters. At the end of this document was the statement, "I have read, understood, and printed a copy of, the above consent form and desire of my own free will to participate in this study," to which PST members could select "no" and decline to participate or "yes" in which case they were directed to the instructions for completing the survey (see Appendix J). PST members were contacted up to three times. Upon PST members' completion of the first survey, *Qualtrics* generated a code for each PST member who completed the PDC-S. *Qualtrics* automatically sent a "thank-you" email to all PST members who completed the first administration and reminded them that a second survey would be sent in approximately two weeks. The researcher monitored *Qualtrics* daily and entered new codes into a codebook, which also included a date on which the researcher would email the PDC-S again for the second administration. The second PDC-S was emailed to all participating PST members 10 days after the completion of administration one. If PST members failed to complete the second administration within six days, a reminder email was sent prompting them to complete the second survey. *Qualtrics* automatically sent another "thank-you" email to all PST members who completed both administrations of the PDC-S and notified them that their

name had been entered into a drawing for a chance to win one of eight $50 gift cards. All participating PST members who completed the PDC-S twice were entered in a drawing to win one of eight $50 gift cards.

Data Analysis

Excel, *Qualtrics,* and *Statistical Package for the Social Sciences* (SPSS) were used to store, manage, and analyze data. Several data analyses were conducted to answer the research questions. All procedures are described below.

Descriptive Statistics. Prior to running data analyses, in order to test the following hypotheses, participants' raw scores were converted into summary scores by the four broad areas on the PDC-S. For response set one ("Is it in place?"), answers of "no" were coded into values of zero and answers of "yes" were coded into values of one. If participants answered all but one item for area one ("training/knowledge"; 8 items) and area four ("task clarification and prompting"; 11 items), their scores were averaged to produce an area score. For area two ("resources, materials & processes"; 15 items) and area three ("performance feedback/consequences & outcomes"; 17 items), participants who completed all but one or two of the items had an average of their responses computed to produce an area score. Based on this percentage for response sets one and three, descriptive statistics were calculated.

Hypothesis 1. *The PDC-S will have strong content validity with the overall percentage of items correctly classified by experts (the factorial validity index; FVI) being $\geq$80%, and the representativeness of the correctly classified items (content validity index; CVI) being $\geq$80% (McGartland-Rubio, 2005; McGartland-Rubio et al., 2003).*

To examine the content validity of the PDC-S, the factorial validity index (FVI) and the content validity index (CVI) were calculated. These indices were calculated by cross-tabulating

the collected data within *Qualtrics*, where the data were stored. The FVI was calculated for each item, the four factors, and the measure as a whole by a two-step process: (1) the number of raters who correctly assigned an item to its designated factor was divided by the total number of experts for that item, and (2) the average of all items was computed to obtain the FVI (McGartland-Rubio, 2005; McGartland-Rubio et al., 2003). Next, the CVI was calculated for each item, the four factors, and the measure as a whole. To calculate the CVI, expert reviewers' responses were dichotomized into two levels with ones and twos being grouped together (Group 1) and threes and fours being grouped together (Group 2). Then, the percentage of responses falling into the Group 2 was calculated to see if it met the 80% minimum (McGartland-Rubio, 2005). Clarity ratings were also obtained and calculated at the item-, factor-, and measure-levels. Clarity ratings were calculated by averaging the ratings from expert reviewers who assigned a particular item to the correct factor. Additional descriptive and qualitative data were gathered from expert reviewers during phase two to assist the researcher with determining the final items for the PDC-S.

Hypothesis 2. *Inter-rater reliability will vary somewhat within PSTs as each member has their own perceptions of the team's processes. However, it is hypothesized that the median correlation among raters within each participating PST will be k* $\geq$*0.60 (DeVellis, 2005). Furthermore, the average percent agreement among raters within each PST will be* $\geq$ *70% (Iwata et al., 2013).* To determine the inter-rater reliability of the PDC-S, both inter-class correlations (ICCs) and Cohen's kappa (k) were calculated for the PST members within a single PST. Kappa divides the number of observed agreements between raters (f_O) minus the number of agreements expected by chance (f_E), by the total number of observations (N) minus the number of agreements expected by chance (f_E):

$$k = (f_O - f_E) / (N - f_E)$$

Kappa can also be thought of as the actual agreements beyond chance divided by the potential agreements beyond chance (DeVellis, 2005).

Percent agreement was calculated for each pair of raters within a participating PST. Percent agreement was calculated by dividing the number of shared responses by the number of overall responses for each of the four areas on the PDC-S. Average percentages of agreement were calculated for each team and area for both response sets one and three.

Hypothesis 3. *Percent agreements will be higher within-groups than between-groups for inter-rater reliability (DeVillis. 2012).* The percent agreements yielded from research question three will be compared across groups to determine the differences between inter-rater reliability within-groups and between-groups.

Hypothesis 4. *The test-retest reliability of the PDC-S will be high, with a correlation coefficient of 0.80 or higher (Webb et al., 2006).* The test-retest reliability for all categories and the PDC-S as a whole will be calculated using both Cronbach's alpha and the Pearson product-moment correlation, looking at the correlation between each PST members' PDC-S results at administration one and administration two (Netemeyer et al., 2003; Webb et al., 2006).

CHAPTER 4

Results

Prior to conducting the aforementioned data analyses to answer the research questions, preliminary data analyses were conducted using SPSS in order to prepare the data files. First, participants' raw scores on the PDC-S were calculated into percentages and/or averages by the four areas, the process of which largely accounted for missing data values. Additionally, descriptive statistics for the PDC-S items were summarized.

Descriptive statistics Table 5 displays the descriptive statistics for all four areas of the PDC-S in regards to the first response set, which asked participants whether or not a particular item was in place within their team (response options: *no* [0] or *yes* [1]). Table 6 displays the descriptive statistics for the third response set, which asked participants "how important" that particular item is (response options: *very important* [1], *somewhat important* [2], or *not important* [3]).

Table 5

Descriptive Statistics for Response Set #1

Area	*n*	Mean	Std. Deviation	Range
Training & Knowledge	152	0.67	0.29	0 – 1.00
Task Clarification & Prompting	154	0.73	0.19	0.07 – 0.93
Resources, Materials & Processes	153	0.72	0.20	0.12 – 0.88
Performance Feedback/ Consequences & Outcomes	150	0.67	0.22	0 – 1.00

Table 6

Descriptive Statistics for Response Set #3

Area	*n*	Mean	Std. Deviation	Range
Training & Knowledge	143	1.40	0.43	1.00 – 3.00
Task Clarification & Prompting	121	1.41	0.31	1.00 – 2.47
Resources, Materials & Processes	119	1.32	0.28	1.00 – 2.22
Performance Feedback/ Consequences & Outcomes	118	1.45	0.30	1.00 – 2.36

Qualitative Analysis

The second response set on the PDC-S asked participants who answered the first response set with "no" to select why that particular activity/action is not currently in place on their team. Table 7 below summarizes the frequency with which each functional category was selected for the four areas on the PDC-S. Percentages are also displayed for each functional category by the four areas. For area one, *Training/Knowledge*, 41.1% of participants selected the functional category "limited knowledge/experience with this activity." For area two, *Task Clarification & Prompting*, two functional categories were identified by 27.6% of participants as to why items in that area were not currently in place within their team: "other needs take priority over this activity/action during team meetings" and "limited knowledge/experience with this activity." The highest number of participants (29.4%) selected the functional category "limited knowledge/experience with this activity" for area 3, *Resources, Materials and Processes*. Finally, 24.7% of participants selected "not able to devote time to support this activity/action" for area four, *Performance Feedback/Consequences and Outcomes*.

Table 7

Frequencies of Functions by Area (Response Set #2)

Functional Category	Area 1	Area 2	Area 3	Area 4	Mean
Other needs take priority over this activity/action during team meetings	27.3%	27.6%	22.8%	22.6%	25.1%
Limited knowledge/experience with this activity	41.1%	27.6%	29.4%	21.0%	29.8%
Lack of communication/direction provided to the team regarding this activity/action	7.8%	21.9%	14.1%	16.3%	15.0%
Limited support/acknowledgement from administration/team for this action	11.4%	6.4%	13.4%	15.4%	11.7%
Unable to devote time to support this activity/action	12.4%	16.5%	20.3%	24.7%	18.5%

Research Question #1

The factorial validity, content validity, and clarity indices were calculated based on the responses from the 17 expert reviewer participants during Phase 2a as well as the four additional expert reviewer participants during Phase 2b in order to assess the content validity of the PDC-S.

Phase 2a. The results for Phase 2a are summarized in the four tables below. Table 8 contains the FVIs, CVIs, and Clarity Ratings for all items within area one, Table 9 displays results for area two, Table 10 for area three, and Table 11 for area four. The FVI for area one, "Training/Knowledge" was the highest of the four areas at 72.37%. Of the expert reviewers who correctly assigned the items within factor number one, a CVI of 98.39% was reported. Clarity was also rated highly for area one at 3.76. The FVI for items under area two, "Task Clarification and Prompting," was 24.67%, whereas the CVI was 96.4%, and the clarity was rated at 3.72. For area three, "Resources, Materials, and Processes," the FVI was 60.04%, the CVI was 97.95%, and the clarity rating was 3.80. As explained above, an additional round of expert review was

conducted to assess the content validity of the revised PDC-S based on the results and feedback of the initial expert panel review (Phase 2a). The FVI for area four, "Performance Feedback and Consequences," was 47.45%, the CVI was 97.38%, and the clarity rating was 3.58. The overall FVI, CVI, and clarity indices were also calculated for the measure as a whole; the FVI was 51.13%, the CVI was 97.55%, and the clarity rating was 3.72. Based on these results, the PDC-S was revised utilizing certain criteria and discussions among the researcher, Project LIST Co-Principal Investigators and Director, and a consultant with expertise in statistical analysis and testing/measurement procedures.

Table 8

Phase 2a: Factor Validity Indices, Content Validity Indices, and Clarity Ratings for Area 1

#	Area 1: Training/Knowledge – Items	FVI[1]	CVI[2]	Clarity[3]
1.	Team members have been trained in the process of analyzing students' behavioral problems.	100.00	93.75	3.63
2.	Team members develop operational definitions for students' behavioral problems.	20.00	100.00	3.33
3.	Team members received training in ways to use data to assess students' behavioral problems.	87.50	100.00	3.71
4.	Team members have been trained on how to establish goals for behavior change.	100.00	100.00	3.88
5.	Team members have been trained to identify and select behavioral interventions that are evidence-based.	93.75	93.33	3.60
6.	Team members consistently link assessment findings/data to intervention selection.	21.43	100.00	4.00
7.	Training is available to assist team members and appropriate staff to accurately implement the selected intervention.	75.00	100.00	4.00
8.	Team members are able to describe "fidelity checks" and when they should be performed.	81.25	100.00	3.91

[1] FVI = Percentage of expert reviewers who correctly classified the item
[2] CVI = The representativeness of the correctly classified items
[3] Clarity = The average rating from participants who assigned the item to its corresponding factor (1=Not clear, 4=Clear)

Table 9

Phase 2a FVIs, CVIs, and Clarity Ratings for Area 2

#	Area 2: Task Clarification & Prompting – Items	FVI	CVI	Clarity
9.	The meeting agenda is communicated in advance of meeting.	46.67	100.00	3.86

10.	Meeting notes are sent out after the meeting.	7.14	100.00	3.00
11.	Each core team member has been assigned a specific role within the team.	62.50	100.00	3.80
12.	Each team member understands their role on the team. If yes, briefly explain your role.	43.75	100.00	3.57
13.	Team members are verbally or electronically (i.e., email or text) reminded to collect or bring data to share at team meetings.	81.25	92.31	3.61
14.	Team members consistently select evidence-based solutions to address student behavior.	6.67	100.00	4.00
15.	The intervention selected to address a student's problem behavior is clear to team members at the end of each team meeting.	57.14	87.50	3.38
16.	Team members consistently select evidence-based solutions to address student behavior.	0.00	N/A	N/A
17.	The team member(s) assigned to implement the intervention is verbally or electronically (i.e., email or text) reminded to complete the implementation of the solution.	86.67	100.00	3.77
18.	The selected intervention(s) are directly linked to data gathered about the problem behavior.	6.67	100.00	4.00
19.	Fidelity checks are completed on time.	6.67	100.00	4.00
20.	Team members are verbally or electronically (i.e., email or text) reminded to complete fidelity checks.	75.00	100.00	3.83
21.	A task analysis of intervention implementation (e.g., a checklist, data sheet with step-by-step instructions) is provided to the team before implementing the intervention.	50.00	100.00	3.75
22.	Team members ask for assistance or clarification prior to beginning intervention implementation.	81.25	100.00	3.85
23.	There is an expectation that team members will bring data to refer to during the team meetings.	31.50	80.00	4.00
24.	The process for conducting team meetings has been clearly established and followed.	12.50	100.00	3.50
25.	All team members participate fully during team meetings.	33.33	80.00	3.40
26.	All team members contribute their ideas/concerns during team meetings and are heard.	6.67	100.00	4.00

Table 10

Phase 2a FVIs, CVIs, and Clarity Ratings for Area Three

#	Area 3: Resources, Materials & Processes	FVI	CVI	Clarity
27.	One or more team members are trained in the collection and analysis of data that are used in decision-making.	6.25	100.00	3.00
28.	Team members consistently collect data that are used in team decision-making.	66.67	100.00	4.00
29.	Team members use collected data to identify the problem.	66.67	90.00	3.80
30.	Team members use collected data to select interventions.	73.33	100.00	3.91
31.	Team members use collected data to monitor intervention effectiveness.	60.00	100.00	3.67
32.	Team membership includes at least one member trained in functional assessment.	37.50	100.00	4.00
33.	On what basis are interventions most often generated? (Participants will be asked to rank order the following	53.33	87.50	3.88

	options 1-4 with 1 = most often generated and 4 = least often generated: Evidence-base, team member experience, workshop idea, data).			
34.	The team generates interventions that are linked to data gathered about the identified goal.	56.25	100.00	3.89
35.	The appropriate materials for intervention implementation are available for staff (e.g., manuals, timers, forms, charts, etc.).	N/A	N/A	N/A
36.	Required materials for assessment and intervention implementation are organized (e.g., materials are kept in a specific location, there is a process for checking materials out, etc.)	100.00	100.00	3.88
37.	Required materials for assessment and intervention implementation are available.	93.75	100.00	3.93
38.	Referred students' progress towards goals is reported at subsequent team meetings.	46.67	100.00	3.86

Table 11

Phase 2a: Factorial Validity Indices, Content Validity Indices and Clarity Ratings for Area 4

#	*Area 4: Performance Feedback & Consequences*	FVI	CVI	Clarity
39.	Defining referred students' behavioral issue(s) during the meeting is viewed as an important first step toward problem resolution.	0.00	N/A	N/A
40.	The use of data to determine intervention decisions is supported and encouraged by all team members.	18.75	100.00	3.50
41.	Team members are recognized and praised when data is used to determine interventions.	87.50	92.86	3.71
42.	Team derived interventions are viewed by staff that implement them as feasible.	13.33	100.00	3.50
43.	A team member is assigned to directly monitor/support staff implementing interventions.	6.25	100.00	3.00
44.	Staff members responsible for implementing interventions receive feedback from team members about intervention implementation.	87.50	100.00	3.79
45.	Team members receive feedback about the implementation or lack of implementation of the intervention from administration.	93.75	93.33	3.79
46.	Team members see the value of completing fidelity checks.	26.67	100.00	3.75
47.	There are consequences for a team member if data are not used to guide intervention decisions.	93.33	92.86	3.60

Phase 2b. As explained above, an additional round of expert review was conducted to assess the content validity of the revised PDC-S based on the results and feedback of the initial expert panel review (Phase 2a). For this phase only the FVI and CVI were calculated for each item, factor, and the overall measure (see Tables 12 – 15). Clarity ratings were not obtained

during this phase. The FVI for area one, "Training/Knowledge" was the highest of the four areas at 77.08% and the CVI was 100%. The FVI for items under area two, "Task Clarification and Prompting," was 67.83% with a CVI of 100%. For area three, "Resources, Materials, and Processes," the FVI was 70.45% and the CVI was 98.49%. The FVI for area four, "Performance Feedback and Consequences," was 42.50% and the CVI was 96.17%. The overall FVI, CVI, and clarity indices were also calculated for the measure as a whole; the FVI was 64.47% and the CVI was 98.77%.

Table 12

Phase 2b: Factorial Validity Indices and Content Validity Indices for Area 1

#	Area 1: Training/Knowledge – Items	FVI[1]	CVI[2]
1.	Team members have been trained on how to establish goals for behavior change.	75.00	100.00
2.	Team members have been trained in the process of analyzing students' behavior problems.	100.00	100.00
3.	Team members received training in data-driven assessment of student behavior problems.	75.00	100.00
4.	One or more team members are trained in the collection and analysis of data that are used in decision-making.	75.00	100.00
5.	Team members have been trained to identify and select behavioral interventions that are evidence-based.	100.00	100.00
6.	Training is available to assist team members and appropriate staff to accurately implement the selected intervention.	75.00	100.00
7.	Team members are able to describe "fidelity of implementation checks" and when they should be performed.[3]	66.67	100.00
8.	Team membership includes at least one member trained in functional assessment.	50.00	100.00

[1] FVI = Percentage of expert reviewers who correctly classified the item
[2] CVI = The representativeness of the correctly classified items
[3] *n*=3 for this item

Table 13

Phase 2b: Factorial Validity Indices and Content Validity Indices for Area 2

#	Area 2: Task Clarification & Prompting – Items	FVI	CVI
9.	Each core team member has been assigned a specific role within the team.	100.00	100.00
10.	Team members understand their role on the team. If yes, briefly explain your role: _____	75.00	100.00

11.	Team members are verbally or electronically (i.e., email or text) reminded to collect or bring data to share at team meetings.	75.00	100.00
12.	By the end of each team meeting, the team ensures all members understand the selected intervention(s) and their assigned roles.	50.00	100.00
13.	A task analysis of intervention implementation (e.g., a checklist containing all intervention components, data sheet with step-by-step instructions) is provided to the team before implementing the intervention.	66.67	100.00
14.	If needed, team members ask for assistance or clarification prior to beginning intervention implementation.	50.00	100.00
15.	Team members assigned to implement the intervention are verbally or electronically (i.e., email or text) reminded to complete the implementation of the intervention.	50.00	100.00
16.	Team members are verbally or electronically (i.e., email or text) reminded to complete fidelity of implementation checks.	75.00	100.00

Table 14

Phase 2b: Factorial Validity Indices and Content Validity Indices for Area 3

#	Area 3: Resources, Materials & Processes	FVI	CVI
17.	The meeting agenda is communicated in advance of the meeting.	25.00	100.00
18.	The process for conducting team meetings has been clearly established and followed.	25.00	100.00
19.	All team members participate fully during team meetings.	25.00	100.00
20.	All team members contribute their ideas/concerns during team meetings and are heard.	25.00	100.00
21.	Team members develop operational definitions for students' behavioral problems.	50.00	100.00
22.	There is an expectation that team members will bring data on individual students to discuss during the team meetings.	25.00	100.00
23.	Team members consistently create behavioral goals for referred students.	75.00	100.00
24.	Team members use collected data to identify the problem.	100.00	100.00
25.	Team members consistently link assessment findings to intervention selection.	100.00	100.00
26.	Team members use collected data to select interventions.	100.00	100.00
27.	Team members consistently select evidence-based solutions to address student problem behavior.	100.00	100.00
28.	Students' progress towards goals is reported at subsequent team meetings.	75.00	100.00
29.	Team members evaluate student progress towards goals.	75.00	100.00
30.	Team members document that interventions were implemented according to the plan.	100.00	100.00
31.	Fidelity of implementation checks are completed regularly.	100.00	100.00
32.	Team members use collected data to monitor intervention effectiveness.	100.00	100.00
33.	It is clear to the team when intervention changes are needed.	75.00	66.67
34.	The appropriate materials for intervention implementation are available for staff (e.g., manuals, timers, forms, charts, etc.).	100.00	100.00
35.	Required materials for assessment and intervention implementation are organized (e.g., materials are kept in a specific location, there is a process for checking materials out, etc.)	100.00	100.00

36.	Required materials for assessment and intervention are available.	100.00	100.00
37.	Team meeting notes are sent out after the meeting.	25.00	100.00
38.	On what basis are interventions most often generated? (Participants will be asked to rank order the following options 1-4 with 1 = most often generated and 4 = least often generated: Evidence-base, team member experience, workshop idea, data).	50.00	100.00

Table 15

Phase 2b: Factorial Validity Indices and Content Validity Indices for Area 4

#	*Area 4: Performance Feedback & Consequences*	FVI	CVI
39.	Team members are recognized and praised when data are used to determine interventions.	100.00	100.00
40.	There are consequences for a team member if data are not used to guide intervention.	100.00	75.00
41.	The use of data to determine intervention decisions is supported and encouraged by all team members.	50.00	100.00
42.	Staff implementing interventions are directly monitored by a team member and provided with support and/or feedback when needed.	25.00	100.00
43.	Staff members responsible for implementing interventions receive feedback from team members about intervention implementation.	50.00	100.00
44.	Team members receive feedback about the implementation or lack of implementation of the intervention from administration.	50.00	100.00
45.	Students referred to the team typically make progress toward the behavioral goal.	0.00	N/A
46.	Funding, materials, and space are adequate to implement intervention plans.	0.00	N/A
47.	The problem-solving team process is effective at helping referred students.	50.00	100.00
48.	The intervention plan developed for referred students is effective at prevention future occurrence of problem behaviors.	0.00	N/A

Table 16 below compares the FVIs and CVIS produced by the first set of expert reviewers (*n*=17) compared with the FVIs and CVIs produced by the second set of experts (*n*=4) after the researchers made revisions to the PDC-S based on results and expert feedback from the first round of experts in phase 2a.

Table 16

Phase 2a and Phase 2b Area- and Measure-Level FVIs and CVIs

	Phase 2a		Phase 2b	
Area	FVI	CVI	FVI	CVI
Training & Knowledge	72.37	98.39	77.08	100.00

Task Clarification & Prompting	24.67	96.46	67.83	100.00
Resources, Materials & Processes	60.04	97.95	70.45	98.49
Performance Feedback/Consequences & Outcomes	47.45	97.38	42.50	96.57
PDC-S Total	51.13	97.55	64.47	98.77

Both the FVIs and CVIs for areas one, two, and three, improved from the first round of expert reviews to the second round. However, the FVI and CVI for area four decreased between phase 2a and phase 2b. For the PDC-S as a whole, the FVI improved from 51.13 during phase 2a to 64.47 during phase 2b, and the CVI improved from 97.55 from phase 2a to 98.77 during phase 2b.

Research Question #2

The inter-rater reliability of the PDC-S was assessed by calculating the percent agreement for each possible pair of raters within a PST. For PSTs with more than two raters, the mean percent agreement across all pairs of team members was used as the teams' overall percent agreement. Percentages of agreement were calculated for each pair of raters, across all four PDC-S areas, and for both response sets one and three. Table 17 below summarizes the percent agreement for each respective team for response set one on the PDC-S.

Table 17

Percent Agreement by Area for Response Set #1

Team	Area 1	Area 2	Area 3	Area 4	Overall
A1	62.50%	86.70%	76.50%	81.80%	76.88%
B1	75.00%	70.00%	64.70%	51.50%	65.30%
C1	37.50%	73.00%	70.60%	81.80%	65.73%
C2	62.50%	77.81%	75.70%	75.73%	72.94%
C3	58.33%	72.06%	60.40%	53.25%	61.01%
C4	60.83%	77.33%	71.38%	56.48%	66.51%
C5	58.33%	74.47%	83.33%	63.63%	69.94%
D1	87.50%	73.30%	41.20%	60.00%	65.50%

D2	12.50%	66.70%	47.10%	54.50%	45.20%
E1	47.22%	57.22%	64.58%	74.22%	60.81%
E2	75.00%	53.30%	64.70%	54.50%	61.88%
E4	37.50%	73.30%	58.90%	72.20%	60.48%
E5	75.00%	68.90%	49.10%	78.77%	67.94%
E7	50.00%	80.00%	43.80%	63.60%	59.35%
E8	37.50%	73.30%	82.40%	81.80%	68.75%
E12	50.00%	69.33%	53.51%	45.46%	54.58%
F1	66.67%	77.77%	70.60%	81.80%	74.21%
G1	33.30%	40.00%	70.60%	63.60%	51.88%
H1	75.00%	60.00%	64.70%	72.70%	68.10%
H2	50.00%	53.30%	52.90%	45.50%	50.43%
H3	65.00%	51.34%	73.45%	49.92%	59.93%
H4	50.00%	55.53%	33.60%	87.27%	56.60%
I1	75.00%	46.70%	58.90%	18.20%	49.70%
K1	50.00%	77.77%	80.43%	81.80%	72.50%
M1	85.00%	86.66%	80.60%	84.51%	84.19%
M2	58.33%	71.87%	65.72%	72.70%	67.16%
N1	75.00%	90.48%	81.18%	90.90%	84.39%
O1	37.50%	53.30%	29.40%	36.40%	39.15%
P1	75.00%	73.37%	52.97%	60.60%	65.49%
P2	100.00%	82.20%	79.80%	84.83%	86.71%
P3	59.59%	67.37%	59.25%	52.71%	59.73%
P4	81.25%	77.33%	67.67%	63.33%	72.40%
Q1	87.50%	86.70%	82.40%	63.60%	80.05%
R1	100.00%	86.70%	82.40%	100.00%	92.28%
R2	64.52%	75.06%	65.90%	62.71%	67.05%
R3	100.00%	73.30%	64.70%	63.60%	75.40%
R4	79.90%	90.30%	88.20%	70.07%	82.12%
S1	100.00%	80.00%	64.70%	90.90%	83.90%
T1	58.33%	75.55%	73.53%	65.15%	68.14%
T2	66.67%	58.41%	52.20%	60.16%	59.36%
U1	87.50%	66.70%	82.40%	72.70%	77.33%
V1	37.50%	46.20%	64.70%	55.60%	51.00%
Total	64.42%	70.25%	65.50%	66.68%	66.71%

Percent agreement for response set one on the PDC-S varied widely across teams and areas, with team averages across areas ranging from 39.15% - 100%. Area one had the lowest average percent agreement of across teams (64.42%), followed by area three (65.50%), then area four (66.68%), and area two yielded the highest average percent agreement across raters for response set one (70.25%). The overall percent agreement across all raters and areas for the PDC-S as a whole was 66.71%.

Table 18

Percent Agreements by Area for Response Set #3

Team	Area 1	Area 2	Area 3	Area 4	Median
A1	62.50%	86.70%	50.00%	72.70%	67.98%
B1	91.67%	33.37%	61.43%	57.57%	61.01%
C1	37.50%	40.00%	38.90%	72.70%	47.28%
C2	60.42%	62.40%	63.03%	54.55%	60.10%
C3	86.30%	64.00%	76.48%	71.51%	74.57%
C4	38.93%	50.00%	63.07%	58.24%	52.56%
C5	22.92%	55.53%	48.53%	51.50%	44.62%
D1[a]	100.00%	80.00%	--	66.70%	82.23%
D2[a]	100.00%	--	45.50%	14.30%	53.27%
E12	51.79%	70.33%	50.13%	58.99%	57.81%
E1	56.95%	42.40%	43.63%	55.73%	49.68%
E2	25.00%	33.30%	55.60%	45.50%	39.85%
E4	87.50%	80.00%	82.40%	55.60%	76.38%
E5	29.17%	40.03%	27.77%	45.43%	35.60%
E7	62.50%	73.30%	30.00%	63.60%	57.35%
F1	50.00%	58.33%	79.67%	66.63%	63.66%
G1	33.30%	33.30%	66.70%	54.50%	46.95%
H1	12.50%	40.00%	72.20%	54.50%	44.80%
H2	25.00%	20.00%	25.00%	42.60%	28.15%
H3	56.25%	48.51%	66.83%	59.13%	57.68%
H4	66.67%	35.70%	44.40%	75.00%	55.44%
K1	50.00%	62.23%	66.67%	72.70%	62.90%
M1	57.50%	78.85%	83.03%	65.46%	71.21%
M2	60.42%	56.48%	51.85%	59.08%	56.96%
N1	54.65%	64.20%	62.08%	67.24%	62.04%
P1	75.00%	33.33%	46.53%	73.07%	56.98%
P2	51.30%	90.47%	92.60%	87.27%	80.41%
P3	33.33%	28.54%	50.21%	37.40%	37.37%
P4	68.75%	81.10%	71.65%	70.32%	72.96%
Q1[a]	75.00%	--	80.00%	90.00%	81.67%
R1	100.00%	86.70%	100.00%	80.00%	91.68%
R2	55.00%	54.71%	43.33%	51.92%	51.24%
R3	100.00%	80.00%	100.00%	81.10%	90.28%
R4	77.97%	61.67%	96.30%	91.67%	81.90%
S1	87.50%	61.50%	78.60%	63.60%	72.80%
T1	75.00%	70.00%	75.00%	69.67%	72.42%
T2	58.93%	46.70%	67.29%	63.02%	58.99%
U1	62.50%	33.30%	55.60%	80.00%	57.85%
V1	87.50%	71.40%	70.60%	60.00%	72.38%
Total	61.21%	56.98%	62.70%	63.09%	61.00%

[a] = Area percent agreement omitted due to missing data

Percent agreement for response set three on the PDC-S varied widely across teams and areas, with team averages across areas ranging from 12.50% - 100%. Area two had the lowest

average percent agreement of across teams (56.98%), followed by area one (61.21%), then area three (62.70%), and area four yielded the highest average percent agreement across raters for response set one (63.09%). The overall percent agreement across all raters and areas for the PDC-S as a whole was 61.00%.

Research Question #3

The difference between within-group percent agreement and between-group percent agreement was analyzed by calculating the mean and standard deviations for both groups by response set. The table below summarizes the results. The mean percent agreement for response one within groups was 68.02% (*SD* = 0.13). Between groups, the mean percent agreement for response set one was 65.95% (*SD* = 0.14). For response set three, the mean percent agreement within groups was 48.95% (*SD* = 0.25). The between-group percent agreement mean was 47.98% (*SD* = 0.24) for response set three.

Table 19

Within-Group Versus Between-Group Differences in Percent Agreement

	Response Set 1		Response Set 3	
Area	*Mean*	*SD*	*Mean*	*SD*
Within Groups	68.02%	0.13	48.95%	0.25
Between Groups	65.95%	0.14	47.89%	0.24

Research Question #4

The results for Phase 3b of this study examined the test-retest reliability of the PDC-S by calculating both Cronbach's alpha (α) and Pearson's product moment coefficients (*r*) for participants' responses on administration one compared to their responses on administration two. Alpha was calculated for all of the four areas on the PDC-S for both response sets one and three.

Alphas calculated for response set one (i.e., "is it in place?") were as follows: area one ("Training/Knowledge"; *n*=108) α=.879; area two ("Task Clarification and Prompting"; *n*=110) α=.813; area three ("Resources, Materials, and Processes"; *n*=108) α=.846; and area four ("Performance Feedback and Consequences"; *n*=105) α=.825. Pearson's coefficients were also calculated for response set one: area one (*n*=108) *r*=.790; area two (*n*=110) *r*=.686; area three (*n*=108) *r*=.733; and area four (*n*=105) *r*=.703.

Alphas were also calculated for each PDC-S area for the third response set (i.e., "how important?") were as follows: area one (*n*=98), α=.676; area two (*n*=95) α=.777; area three (*n*=94) α=.805; and area four (*n*=86) α=.803. Pearson's reliability coefficients were also calculated for response set three and were as follows: area one (*n*=98) *r*=.549; area two (*n*=95), *r*=.637; area three (*n*=94), *r*=.686; and area four (*n*=86) *r*=.671.

Table 20

Cronbach Alphas and Pearson Correlation Coefficients by Area

Area	Response Set #1		Response Set #3	
	α	*r*	α	*r*
Training & Knowledge	.879	.790	.676	.549
Task Clarification & Prompting	.813	.686	.777	.637
Resources, Materials & Processes	.846	.733	.805	.686
Performance Feedback/Consequences & Outcomes	.825	.703	.803	.671

CHAPTER 5

Discussion

This study was conducted to adapt a pre-existing measure currently used in the field of Organizational Business Management that employs principles of Applied Behavioral Analysis and Functional Behavioral Analysis to increase employee productivity and overall functioning within various businesses/organizations. Schools, as well as school-based PSTs, are organizations that share many commonalities with more traditional "businesses." In addition to adapting the *Performance Diagnostic Checklist* into the *Performance Diagnostic Checklist for Schools* for application with school teams, the current study examined initial psychometric properties of the PDC-S including the content validity, inter-rater reliability, and test-retest reliability.

Qualitative Analysis

The functional categories on the PDC-S, or response set two, are intended to assist in the identification of reasons why certain practices are not currently taking place within a PST. Individuals were only allowed to select one functional category for each item that they responded with "no" to response set one. Results showed that for all areas with the exception of area four, "Performance Feedback/Consequences and Outcomes," the most frequently cited functional category was *limited knowledge/experience with this activity*. This finding is consistent with previous literature (Conoley et al., 2009; Reinke et al., 2011) citing both lack of knowledge and training as a barrier in selecting and implementing EBIs. In the current study, the functional category, *other needs take priority over this activity/action during team meetings*, was a close second for areas one through four. Williamson and McLeskey (2011) found that it is common

for teams to get side-tracked during meetings and spend a large portion of their meeting time on tangential issues not directly related to the issue(s) at hand.

The frequency with which specific functional categories were selected for certain areas was not surprising given their nature. For example, for area one, "Training/Knowledge," the functional category of *limited knowledge/experience with this* activity was the most frequently selected reason for why a certain task or activity was not currently in place. Furthermore, for area two, "Task Clarification and Prompting," the functional category, *lack of communication/direction provided to the team regarding this activity*, was most frequently selected. In other words, PST members cited plausible reasons for why a certain task or activity was not currently in place for their respective teams.

The literature lists as one of the most commonly cited reasons for failure to implement all steps of the problem-solving process, including the selection and implementation of EBIs, is due to a lack of time and/or resources (including staff ability; Lopes et al., 2004; Reinke et al., 2011). Therefore, it is expected that the functional category *unable to devote time to support this activity/action* would receive between 12.4% and 24.7% of the endorsements, depending on the area.

Another barrier cited in the literature to effective problem-solving teaming is failure to fit with the school mission (Adelman & Taylor, 2003; Walker, 2004). Although the functional category *limited support/acknowledgement from administration/team for this action* received the lowest percentage of endorsements for all areas with the exception of area one, it still yielded percentages ranging from 6.4% - 15.4%. Furthermore, it is important to remember that participants were only allowed to select one functional category per item. Thus, it is possible that there was more than one reason as to why a team member thought that their team was not

currently fulfilling a specific task/activity, but they were forced to select the answer that best explained their selection.

Hypothesis 1. *The PDC-S will have strong content validity with the overall percentage of items correctly classified by experts (the factorial validity index; FVI) being ≥80%, and the representativeness of the correctly classified items (content validity index; CVI) being ≥80% (McGartland-Rubio, 2005).*

The FVIs varied widely across factors with none of the factor or overall FVIs exceeding 80% during either Phase 2a or 2b, failing to support the first part of the hypothesis. There are several possible reasons for why the PDC-S did not meet the 80% benchmark for the FVI. Feedback from expert reviewers indicated two main issues with the items on the PDC-S. The first issue expert reviewers noted was that certain items could fall into two categories or factors. The second issue was that there seemed to be additional factors tapped with the items presented. Specifically, several expert reviewers noted that the items seemed to also be tapping into best practices and various steps in the problem-solving team process. The improvement in FVIs between Phase 2a and 2b for areas one through three suggest that, overall, the revisions made based on the results and feedback during Phase 2a were beneficial. However, there were still some items and areas that produced low FVIs. Some of this was due to an inconsistency in ratings between experts in Phases 2a and Phase 2b. For example, the experts in Phase 2a may have indicated that a certain item was most representative of a different area than it was assigned to, so that item was moved to the new area for Phase 2b, but the experts in Phase 2b indicated that that item is most representative of the original area which it was assigned to. Thus, some items with a FVI of below 80% were retained after Phase 2b but moved back to the area they were originally signed to. Additional items that did not meet the FVI criterion of ≥80% were

kept because the information yielded from those items were important to the overall intent and treatment utility of the PDC-S.

The FVI for area four in particular ("Performance Feedback/Consequences & Outcomes") was low for both Phase 2a and Phase 2b (47.50% and 42.50%, respectively). Since having an unclear factor(s) can cause low FVIs, it is possible that the specificity of area four was the cause of the consistently low FVIs across both phases (McGartland-Rubio, 2005; McGartland-Rubio et al., 2003). In order to address this issue, the title of area four was changed from "Performance Feedback & Consequences" as it was for Phases 2a and 2b to "Performance Feedback/Consequences & Outcomes" for the administration to the developmental sample in Phase 3. Some of the items related to student outcomes received very low FVIs from the expert reviewers since they did not really fit the original area four of "Performance Feedback & Consequences." Thus by adding the "Outcomes" portion to area four, it should improve the FVI if a third expert panel review was to be conducted.

The content validity indices were strong for all factors across both Phases 2a and 2b ($\geq$96.46%), indicating that, for the expert reviewers who correctly assigned items to their designated factors, participants believed that those items were highly representative of the factor to which they were assigned. Thus, part two of the hypothesis was supported. The clarity indices were also high across all items and areas for Phase 2a. Having CVIs that are consistently higher than the FVIs, in addition to high clarity ratings, indicates that the wording of the items is clear and representative of the respective factors, but the actual factor that the item is supposed to be representing may not be appropriate and/or exclusive from other possible factors (McGartland-Rubio, 2005; McGartland-Rubio et al., 2003). Thus, significant changes did not need to be made to the wording of the items based on the ratings from the panel of expert

reviewers. However, revision of the four areas may be warranted in order to create clearer factors, producing higher FVIs, and thus yielding stronger content validity across the PDC-S. One potential option for changing the areas would be to align the items of the PDC-S with the DORA II categories used during team based direct observations (Algozzine et al., 2011).

Hypothesis 2. *Inter-rater reliability will vary somewhat within PSTs as each member has his or her own perception of the team's processes. However, it is hypothesized that the average percent agreement among raters within each PST will be ≥ 70% (Iwata et al., 2013).*

Inter-rater reliability, or agreement, was calculated using percent agreement among pairs of raters within each PST for both response sets one and three of the PDC-S. An overall percent agreement for each of the four areas was also calculated. For response set one, results indicated that area two, "Task Clarification and Prompting," had the strongest percent agreement, meeting the 70% standard. However, area two also possessed the lowest percent agreement for response set three. Thus, although participants had strong agreement regarding whether or not an item under "Task Clarification and Prompting" was currently in place within their team, there was low agreement as to whether or not that is important to the overall PST process.

The literature tells us that one barrier to effective teaming is philosophical differences between team members regarding the consultation process (Conoley et al., 2009; Nellis, 2012). Specifically, general education teachers tend to see less value in consultation and sometimes feel as though their opinion is not valued within the PST setting and process (Williamson & McLeskey, 2011). Given that the largest respondent group in the current study was general education teachers (27.1%), variance in team member perspectives is expected, especially in regards to response set three, which asked participants to indicate "how important" certain tasks within the problem-solving process are to them.

For response set three, area four, "Performance Feedback/Consequences & Outcomes," yielded the strongest percent agreement among PST members. Overall, response one had a stronger total percent agreement across areas and PSTs than response set three did. This is fitting given that response set one measures whether or not a task or activity is currently in place within a team; this type of rating is more objective and thus warrants greater agreement among raters than asking "how important" said task or activity is (i.e., response set three). It is to be expected that there would be variation in PST members' opinions regarding the importance of specific team tasks and activities given the research supporting differences in perceptions regarding the consultation process (Conoley et al., 2009). Therefore, inclusion of this set of analyses may not be the most useful analysis of the strength of this instrument.

Hypothesis 3. *Correlations will by stronger within PSTs than between PSTs for inter-rater reliability (DeVillis, 2005).* Although the within-group percent agreements were higher than the between-group percent agreements, the difference was not significant. Response set one only had two possible answers, allowing for little variation between raters. Furthermore, the majority of raters across teams and items selected 1, or "yes", for their answer to response set one. With response set three, the mean percent agreement was also slightly higher within groups than between groups, but not significantly. This finding is less surprising for response set three than response set one given that the raters, or team members, were asked to rate their opinion regarding the importance of an item, rather than to assert whether or not a certain practice was currently in place. The literature tells us that disagreement between team members regarding student referral concerns and the importance of the consultation process is common (Nellis, 2012; Williamson & McLeskey, 2011). Overall, these results indicate that the agreement found between PST members for research question three are due at least in part to chance.

Hypothesis 4. *The test-retest reliability of the PDC-S will be high, with a correlation coefficient of 0.80 or higher (Webb et al., 2006).* Both Cronbach alphas and Pearson's product-moment coefficients were calculated for the four areas on the PDC-S for both response sets one and three. All four areas produced Cronbach alphas of $\geq$.80 for response set one, suggesting strong test-retest reliability for the PDC-S (Webb et al., 2006). Overall, alphas and coefficients for response set one were stronger than those for response set three, indicating that the measurement of whether or not something is in place on a team is more stable than team members' beliefs regarding how important a particular action or activity is to team functioning. The discrepancy between correlation coefficients on response set one versus response set three could be attributed to several things. Firstly, response set one measured a more objective construct than did response set three. Furthermore, response set one asked participants to rate a more stable construct than did response set three. That is, participants are more likely to accurately recall whether or not their team completes a certain task and/or activity than they are how important they deemed that task; the value an individual places on the importance of something can change over time and depending on several factors, including an individual's mood on that particular day. Participants' perceptions regarding the importance of an item was also suspect to change over the lapse between administration one and two as they were provided time to reflect on the items.

Limitations

The small sample size (n=4) for Phase 2b (the second expert panel review) is a limitation to this study. Best practices in scale development tell us that a minimum of six panel members is recommended, while 10-15 is optimal (Netemeyer et al., 2003). Having a sample of four expert reviewers on the second panel was also unfortunate since four is an equal number, and decisions

regarding item and scale revisions were based upon a "majority rules" principal (e.g., if more than 50% of experts assigned an item to an area different than its designated area, the item was moved to the area that the experts assigned it to). A panel of four experts also made it difficult to meet the hypothesis since it required all four experts to assign the item to the correct area. Even if one expert reviewer selected the wrong area, that dropped the FVI to 75%, not meeting the 80% criterion. Thus, some items were retained although they did not meet the 80% FVI criterion because either (a) there was inconsistency in FVI ratings between Phase 2a and Phase 2b (e.g., results from Phase 2a indicate that a particular item fit in one area, but after the item was moved to the assigned area for Phase 2b and (b) they yielded important information for intervention decision-making although they did not concisely fit into one particular area.

The composition, size, and meeting frequency of the numerous school teams who participated in this study varied greatly. This finding is likely reflective of the greater population, however it makes it difficult to ascertain what team factor(s) contribute to either the high or low reliabilities found in this study. Another limitation of this study is the homogeneity of the sample. Although the sample was representative of the school personnel for the region (i.e., largely white females), it is likely not representative of the greater nationwide population of school personnel demographics (National Center for Education Statistics, 2012).

One of the limitations present in the current study is the calculation of areas scores based on participants' raw PDC-S scores. Raw scores were calculated into percentages, allowing the researcher to largely account for the missing values present for response set one. However, that procedure assumes that the participant was likely to score the missing item similarly to their other response scores. Ideally, missing values would have been accounted for by returning the

survey to the participant to complete any missed items. Since the PDC-S was administered online, and the responses were kept confidential, this was not a viable option.

Furthermore, due to the nature of the data collected (i.e., the presence of several "constant variables" and missing values), it was not possible to calculate Cohen's kappa as an estimate of inter-rater reliability. While the percent agreement values provide a decent estimate of the inter-rater reliability of the PDC-S, kappa controls for change agreement, making it a stronger statistic (Cohen, 1960).

Due to the lack of an adequate sample size, the current study did not include a factor analysis on the PDC-S. Factor analyses allow researchers to ascertain what latent variables are being tapped into within each construct measured on a scale (DeVellis, 2012). This is an important step of scale development; however, a minimum sample size of 300 is typically necessary to conduct a factor analysis, which was beyond the scope of the current study.

Future Research

The push for schools to utilize evidence-based practices and interventions for students experiencing social, emotional, and behavioral challenges is constant. However, few resources are available to schools that provide the necessary training and assistance to help them be successful in realizing this goal. The current research involved development of an instrument that can identify school teams' barriers to selecting and implementing EBIs. The next step in research would be to, based on the barriers identified by the *Performance Diagnostic Checklist for Schools*, provide effective training models to assist schools in removing identified barriers in order to promote PSTs' uptake of EBIs. For example, the functional category of *limited knowledge/experience with this activity* was most frequently cited as the reason for why a particular task or activity was not currently in place on a PST; thus, future research projects

should focus on developing intervention and/or training modules that increase PSTs' knowledge and experience with important team problem-solving processes. Project LIST, the IES grant within which this study is embedded, is currently working on this task.

In order to most effectively accomplish this goal, additional development and assessment of the PDC-S's psychometric properties would be beneficial. Specifically, conducting a factor analysis of the PDC-S would be beneficial to the ongoing development and validation of the tool. The PDC-S data from Project LIST participants, in combination with the data from the current study, would provide a large enough sample size (i.e., $\geq$300) to conduct a factor analysis to get a better estimate of the content validity of the PDC-S (DeVellis, 2012). Conducting a factor analysis would also help to finalize the items on the PDC-S, which would be beneficial given that not all items or areas met the recommended FVI criterion of $\geq$80%.

In addition to or rather than a factor analysis, a third expert panel review could be conducted in order to further fine-tune the PDC-S item pool. This would provide an assessment of the FVIs and CVIs of the items after they were moved into different areas based upon the results from Phase 2b of the current study. Furthermore, if an additional sample of 10-15 expert reviewers was obtained for this additional panel review, the results would be more useful compared to the results from Phase 2a with a sample of four experts.

Results from the current study indicate low inter-rater reliability among PST members. However, the literature tells us that disagreement among team members is present regarding many issues within the PST process (Nellis, 2012; Williamson & McLeskey, 2011). Thus, it may not be reasonable to expect high inter-rater agreement among team members on the PDC-S. Given this, it is important to consider what the best method would be for gathering and interpreting data for the PDC-S. Austin and colleagues (1999), as well as Carr and colleagues

(2013) have only one rater/observer complete the PDC and PDC-HS, then use the results from that individual to inform decision-making about appropriate interventions. It might be more useful to have the PST leader complete the PDC-S and use the results from their responses, rather than aggregating data collected from all PST members. However, doing this procedure would not allow for input and review of all core team members views on how the team is functioning, which is one benefit of having all team members complete the PDC-S. Weighing the advantages and disadvantages of each type of administration is needed. Further evaluation of the data that has been collected by Project LIST would allow examination of this issue directly by evaluating the PDC-S outcomes with all team members' scores and then compare to the score of the PST leader. This information would provide a way to determine what is gained or lost with each type of administration.

Additionally, it would be both interesting and important to assess the concurrent validity of the PDC-S. The concurrent validity could be assessed by comparing PST members' responses on the PDC-S with trained observers' scores on the DORA II (Aglozzine et al., 2011). Both measures assess similar team processes, but the DORA II provides a more objective assessment of team processes compared to the subjectivity and bias that can be inherent in self-reports (DeVellis, 2012). The discrepancies between scores on the two measures could also be quite informative regarding areas in which to focus intervention efforts.

Summary

The purpose of this study was to adapt the original PDC into the PDC-S, a tool appropriate for application with school-based PSTs, and conduct an initial assessment of the tool's psychometric properties. This research is the first step in establishing the validity and reliability for the PDC-S. Results yielded from this study found that some areas of the PDC-S

contain strong content validity, while other areas need revisions in order to obtain similar factorial validity indices. The inter-rater reliability, as assessed by percent agreement between raters, was just below adequate. Test-retest reliability for response set one was strong. Although reliability coefficients for response set three did not meet the hypothesis, they were still considered moderate. Overall, the PDC-S possesses some strong psychometric properties. This study yielded critical information for further research and additional scale revisions to strengthen the psychometric properties in need of improvement. The PDC-S has promise to become a psychometrically sound and useful tool in conducting functional assessments of PSTs in order to identify areas for training, with the ultimate goal of increasing teams' uptake of EBIs.

References

Adelman, H. S., & Taylor, L. (2003). Rethinking school psychology (commentary on public health framework series). *Journal of School Psychology, 41*, 83-90.

Alberto, P.A., & Troutman, A.C. (2009). Applied behavior analysis for teachers (8th ed.). Upper Saddle River, NJ: Pearson.

Algozzine, B., Newton, J.S., Horner, R.H., Todd, A.W., & Algozzine, K. (2011). Development and technical characteristics of a team decision-making assessment tool: Decision, observation, recording, and analysis (DORA II). *Journal of Psychoeducational Assessment, xx*(x), 1-13.

Allen, S.J., & Graden, J.L. (2002). Best practices in collaborative problem solving for intervention design. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 565-582). Bethesda, MD: National Association of School Psychologists.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for educational and psychological testing.* Washington DC: American Educational Research Association.

American Psychological Association Task Force on Evidence-Based Practice for Children and Adolescents. (2008). *Disseminating evidence-based practice for children and adolescents: A systems approach to enhancing care*. Washington, DC: American Psychological Association.

Asmus, J.M., Vollmer, T.R., & Borrero, J.C. (2002). Functional behavioral assessment: A school based model. *Education and Treatment of Children, 25*(1), 69-90.

Austin, J. (2000). Performance analysis and performance diagnostics. In J. Austin & J. E. Carr (Eds.), *Handbook of applied behavior analysis* (pp. 321–349). Reno, NV: Context

Press.

Austin, J., Carr, J. E., & Agnew, J. (1999). The need for assessment of maintaining variables in OBM. *Journal of Organizational Behavior Management, 19*, 59–87.

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

Bahr, M.W. & Kovaleski, J.F. (2006). The need for problem solving teams: Introduction to the special issue. *Remedial and Special Education, 27*(1), 2-5.

Barton-Arwood, S.M., Wehby, J.H., Gunter, P.L., & Lane, K.L. (2003). Functional behavior assessment rating scales: Interrater reliability with students with emotional or behavioral disorders. *Behavioral Disorders, 28*(4), 386-400.

Betz, A.M., & Fisher, W.W. (2011). Functional analysis: History and methods. In W.W. Fisher, C.C. Piazza, & H.S. Roane (Eds.), *Handbook of applied behavior analysis* (pp. 206-225). New York, NY: Guildford Press.

Bijou, S.W., Peterson, R.F., & Ault, M.H. (1968). A method to integrate descriptive and experimental field studies at the level of data and empirical concepts. *Journal of Applied Behavior Analysis, 1,* 175-191.

Brown-Chidsey, R., & Andren, K.J. (2013). *Assessment for intervention: A problem-solving approach* (2nd ed.). New York, NY: Guildford Press.

Burns, M.K. (1999). Effectiveness of special education personnel in the intervention assistance team model. *The Journal of Educational Research, 92*, 354-356.

Burns, M.K., Peters, R., & Noell, G.H. (2008a). Using performance feedback to enhance implementation fidelity of the problem-solving team process. *Journal of School Psychology, 46*, 537-550.

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

Burns, M.K., Wiley, H.I., & Viglietta, E. (2008b). Best practices in implementing effective problem-solving teams. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V* (pp.1633-1644). Bethesda, MD: National Association of School Psychologists.

Carr, J.E., Wilder, D.A., Majdalany, L., Mathisen, D., & Strain, L.A. (2013). An assessment-based solution to a human-service employee performance problem: An initial evaluation of the performance diagnostic checklist – human services. *Behavior Analysis in Practice, 6*(1), 16-32.

Carr, E.G. (1977). The motivation of self-injurious behavior: A review of some hypotheses. *Psychological Bulletin, 84,* 800-816.

Chaflant, J.C. VanDusen Pysh, M., & Moultrie, R. (1979). Teacher assistance teams: A model for within-building problem solving. *Learning Disability Quarterly, 2*(3), 85-96.

Cohen, J.A. (1960). A coefficient of agreement for nominal scales. Educational and *Psychological Measurement, 20,* 37-46.

Comrey, A.L. (1973). *A first course in factor analysis.* New York: Academic Press.

Conoley, C.W., Conoley, J.C., & Reese, R.J. (2009). Changing a field of change. *Journal of Educational and Psychological Consultation, 19,* 236-247.

Cook, D.A. & Beckman, T.J. (2006). Current concepts in validity and reliability for psychometric instruments: Theory and application. *The American Journal of Medicine, 119*(2), 166.e7-166.e16.

Coplan, R.J., & Rubin, K.H. (1998). Exploring and assessing nonsocial play in the preschool: The development and validation of the preschool play behavior scale. *Social Development, 7*(1), 72-91.

Costello, E. J., Mustillo, S., Keeler, G., & Angold, A. (2004). Prevalence of psychiatric disorders in childhood and adolescence. In B. L. Levin, J. Petrila, & K. D. Hennessy (Eds.), *Mental health services: A public health perspective* (pp. 111–128). New York, NY: Oxford University Press.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart & Winston.

DeVellis, R.F. (2005). Inter-rater reliability. In K. Kempf-Leonard (Ed.), *Encyclopedia of social measurement* (Vol. 2, pp. 317-322). San Diego: Elsevier.

DeVellis, R.F. (2012). *Scale development: Theories and applications* (3rd ed.). Thousand Oaks, CA: SAGE Publications.

Doll, B., Gaack, K., Kosse, S., Osterloh, M., & Siemers, E. (2005). The dilemma of pragmatics: Why schools don't use quality team consultation practices. *Journal of Educational and Psychological Consultation, 16*, 127–155.

DuPaul, G. J. (2003). Commentary: Bridging the gap between research and practice. *School Psychology Review, 32*, 178–180.

Durand, V.M., & Crimmins, D.B. (1988). Identifying the variables maintaining self-injurious behavior. *Journal of Autism and Developmental Disorders, 18*, 99-117.

Eliasson, A.C., Krumlinde-Sundholm, L., Rosblad, B., Beckung, E., Arner, M., Ohrvall, A.M., & Rosenbaum, P. (2006). The manual ability classification system (MACS) for

children with cerebral palsy: Scale development and evidence of validity and reliability. *Developmental Medicine & Child Neurology, 48,* 549-554.

Farmer, E. M., Burns, B. J., Phillips, S. D., Angold, A., & Costello, E. J. (2003). Pathways into and through mental health services for children and adolescents. *Psychiatric Services, 54*(1), 60–66.

Feeney-Kettler, K.A. (2009). Early identification of preschool students at risk for emotional and behavioral disorders: Development and validation of a parent-teacher screener. (Doctoral dissertation). Retrieved from Proquest.

Feeney-Kettler, K.A., Kratochwill, T.R., & Kettler, R.J. (2011). Identification of preschool children at risk for emotional and behavioral disorders: Development and validation of a universal screening system. *Journal of School Psychology, 49,* 197-216.

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, D., Mock, D., Morgan, P.L., & Young, C.L. (2003). Responsiveness-to-intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research and Practice, 18,* 157-171.

Gaumer Erickson, A., Noonan, P., Supon Carter, K., McGurn, L, & Purifoy, E. (2014). The team functioning scale: Evaluating and improving effectiveness of school teams. *International Journal of Educational Research, 69*, 1-11.

Graden, J.L., Casey, A., Christenson, S.L. (1985). Implementing a prereferral intervention system: Part I. The model. *Exceptional Children, 51*(5), 377-384.

Gresham, F.M., Watson, T. S., & Skinner, C.H. (2001). Functional behavioral

assessment: Principles, procedures, and future directions. *School Psychology Review, 30*(2), 156-172.

Glisson, C. (2007). Assessing and changing organizational culture and climate for effective services. *Research on Social Work Practice, 17*, 736-747.

Glisson, C., Landsverk, J., Schoenwald, S.K., Kelleher, K., Hoagwood, K.E., Mayberg, S., & Green, P. (2008). Assessing the organizational social context (OSC) of mental health services: Implications for implementation research and practice. *Administration and Policy in Mental Health and Mental Health Services Research, 35*(1), 98-113.

Glisson, C., & Schoenwald, S.K. (2005). The ARC organizational and community intervention strategy for implementing evidence-based children's mental health treatments. *Mental Health Services Research, 7*(4), 243-259.

Hanley, G.P., Iwata, B.A., & McCord, B.E. (2003). Functional analysis of problem behavior: A review. *Journal of Applied Behavior Analysis, 36*(2), 147-185.

Haynes, S.N., Richard, D.C.S., & Kubany, E.S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment, 7*(3), 238-247.

Hoagwood, K., Burns, B.J., Kiser, L., Ringeisen, H., & Schoenwald, S.K. (2001). Evidence-based practice in child and adolescent mental health services. *Psychiatric Services, 52*(9), 1179-1189.

Individuals with Disabilities Education Improvement Act of 2004, 20 U.S.C. § 1400 *et seq*.

Iverson, A.M. (2002). Best practices in problem-solving team structure and process. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 657-669). Bethesda, MD: National Association of School Psychologists.

Iwata, B.A., DeLeon, I.G., & Roscoe, E.M. (2013). Reliability and validity of the functional analysis screening tool. *Journal of Applied Behavior Analysis, 46*, 271-284.

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1982/1994). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis, 27*, 197-209. (Reprinted from *Analysis and Intervention in Developmental Disabilities, 2*, 3-20, 1982.)

Iwata, B. A., Pace, G. M., Cowdery, G. E., & Miltenberger, R. G. (1994). What makes extinction work: An analysis of procedural form and function. *Journal of Applied Behavior Analysis, 27*, 131-144.

Jimerson, S.R., Burns, M.K., & VanDerHeyden, A.M. (2007). *The handbook of response to intervention: The science and practice of assessment and intervention*. New York, NY: Springer.

Kratochwill, T.R. (2007). Preparing psychologists for evidence-based school practices: Lessons learned and challenges ahead. *American Psychologist, 62*, 829-844.

Kratochwill, T.R. (2008). School-based problem-solving consultation: Applications in prevention and intervention systems. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V* (pp. 1673-1688). Bethesda, MD: National Association of School Psychologists.

Kratochwill, T.R., & Asmus, J. (2011). System-level analysis of evidence-based intervention implementation by problem-solving teams. Special Education Policy, Finance, and Systems section Goal 2. University of Wisconsin – Madison: Madison, WI.

Kratochwill, T.R., & Shernoff, E.S. (2004). Evidence-based practice: Promoting evidence-based interventions in school psychology. *School Psychology Quarterly, 18*(4), 389-408.

Kratochwill, T. R., & Stoiber, K. C. (2000). Diversifying theory and practice: Expanding boundaries of empirically supported interventions in schools. *Journal of School Psychology, 38,* 349-358.

Lopes, J., Monteiro, I., & Sil, V. (2004). Teachers' perceptions about teaching problem students in regular classrooms. *Education and the Treatment of Children, 27,* 394–419.

Masi, R., & Cooper, J.L. (2006). *Children's mental health: Facts for policymakers.* New York: National Center for Children in Poverty, Mailman School of Public Health, Columbia University.

Matson, J.L., Bamburg, J.W., & Cherry, K.E. (1999). A validity study on the questions about behavioral function (QABF) scale: Predicting treatment success for self-injury, aggression, and stereotypes. *Research in Developmental Disabilities, 20*(2), 163-176.

Matson, J.L., Tureck, K., & Rieske, R. (2012). The questions about behavioral function (QABF): Current status as a method of functional assessment. *Research in Developmental Disabilities, 33,* 630-634.

Matson, J.L. & Vollmer, T.R. (1995). *User's guide: Questions about behavioral function (QABF).* Baton Rouge, LA: Scientific Publishers.

McGartland-Rubio, D. (2005). Content validity. In K. Kempf-Leonard (Ed.), *Encyclopedia of social measurement* (Vol. 2, pp. 317-322). San Diego: Elsevier.

McGartland-Rubio, D., Berg-Weger, M., Tebb, S.S., Lee, E.S., & Rauch, S. (2003). Objectifying content validity: Conducting a content validity study in social work research. *Social Work Research, 27*(2), 94-104.

Messick, S. (1993). Validity. In R.L. Linn (Ed.), *Educational Measurement* (2nd ed., pp. 105-

146). Phoenix, AZ: American Council on Education and the Oryx Press.

National Center for Education Statistics (2012). Number and percentage distribution of teachers in public and private elementary are secondary schools [Table 209.10]. Retrieved from: http://nces.ed.gov/programs/digest/d13/tables/dt13_209.10.asp

National Research Council and Institute of Medicine. (2009). Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities. Committee on Prevention of Mental Disorders and Substance Abuse Among Children, Youth and Young Adults: Research Advances and Promising Interventions. M. E. O'Connell, T. Boat, and K.E. Warner, Eds. *Board on Children, Youth, and Families, Division of Behavioral and Social Sciences and Education.* Washington, DC: The National Academies Press.

Nellis, L.M. (2012). Maximizing the effectiveness of building teams in response to intervention implementation. *Psychology in the Schools, 49*(3), 245-256.

Netemeyer, R.G., Bearden, W.O., & Sharma, S. (2003). *Scaling procedures: Issues and applications*. Thousand Oaks, CA: Sage.

No Child Left Behind Act of 2001, P.L. 107-110, 115 Stat. 1425.

Nunnally, J., & Bernstein, I.H. (1994). *Psychometric theory.* (3rd ed.) New York: McGraw-Hill.

O'Neill, R.E., Horner, R.H., Albin, R.W., Sprague, J.R., Storey, K., & Newton, J.S. (1997). *Functional assessment and program development for problem behavior: A practical handbook.* Pacific Grove, CA: Brooks/Cole Publishing.

Paclawskyj, T.R., Matson, J.L., Rush, K.S., Smalls, Y. & Vollmer, T.R. (2000). Questions about behavioral function (QABF): A behavioral checklist for functional assessment of aberrant behavior. *Research in Developmental Disabilities, 21*, 223-229.

Pampino, R.J., Heering, P.W., Wilder, D.A., Barton, C.G., & Burson, L.M. (2003). The use of the performance diagnostic checklist to guide intervention selection in an independently owned coffee shop. *Journal of Organizational Behavior Management, 23*(2/3), 5-19.

Prasse, D.P. (2006). Legal supports for problem-solving systems. *Remedial and Special Education, 27*(1), 7-15.

Reinke, W.M., Stormont, M., Herman, K.C., Puri, R., Goel, N. (2011). Supporting children's mental health needs in schools: Teacher perceptions of needs, roles, and barriers. *School Psychology Quarterly*, *26*(1), 1-13.

Rodriguez, M., Wilder, D.A., Therrien, K., Byron, W., Miranti, R., Daratany, K., Salume, G., Baranovsky, G., & Rodriguez, M. (2006). Use of the performance diagnostic checklist to select an intervention designed to increase the offering of promotional stamps at two sites of a restaurant franchise. *Journal of Organizational Behavior Management, 25*(3), 17-35

Rohn, D., Austin, J., & Lutrey, S.M. (2003). Using feedback and performance accountability to decrease cash register shortages. *Journal of Organizational Behavior Management, 22(*1), 33-46.

Rosenfield, S. A., & Gravois, T. A. (1996). *Instructional consultation teams: Collaborating for change.* New York, NY: Guilford Press.

Sattler, J.M. (2008). *Assessment of children: Cognitive applications* (5th ed.). Jerome M. Sattler, Publisher, Inc.: La Mesa, CA.

Schaughency, E., & Ervin, R. (2006). Building capacity to implement and sustain effective practices to better serve children. *School Psychology Review, 35,* 155–166.

Schoenwald, S., & Hoagwood, K. (2001). Effectiveness, transportiability, and dissemination

of interventions: What matters when? *Psychiatric Services, 52,* 1190-1197.

Sheridan, S. M., & Kratochwill, T. R. (2008). *Conjoint behavioral consultation: Promoting family-school connections and interventions* (2nd ed.). New York, NY: Springer.

Shrout, P. E. & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420-428.

Sugai, G., Horner, R. H., Dunlap, G., Hieneman, M., Lewis, T. J., Nelson, C. M., Scott, T., Liaupsin, C., Sailor, W., Turnbull, A. P., Turnbull, H. R., III, Wickham, D., Reuf, M., & Wilcox, B. (2000a). Applying positive behavioral support and functional behavioral assessment in schools. *Journal of Positive Behavioral Interventions and Support, 2,* 131–143.

Sugai, G., Lewis-Palmer, T., & Hagan-Burke, S. (2000b). Overview of the functional behavioral assessment process. *Exceptionality, 8*(3), 149-160.

Steege, M. W., & Pratt, J. L. (2013). Functional behavioral assessment: The cornerstone of effective problem solving. In R. Brown-Chidsey & K.J. Andren (Eds.), *Assessment for intervention: A problem-solving approach* (2nd ed, pp. 125-143). New York, NY: Guilford Press.

Steege, M. W., & Watson, T. S. (2009). *Conducting school-based functional behavioral assessments: A practitioner's guide* (2nd ed.). New York, NY: Guildford Press.

Tilly, W. D., III. (2008). The evolution of school psychology to science-based practice. In A. Thomas & J.P. Grimes (Eds.), *Best practices in school psychology V* (pp. 18-32). Bethesda, MD: National Association of School Psychologists.

Walker, H. M. (2004). Commentary: Use of evidence-based interventions in schools: Where

we've been, where we are, and where we need to go. *School Psychology Review*, *33*(3), 398-407.

Webb, N. M., Shavelson, R. J., & Haertel, E. H. (2006). Reliability coefficients and generalizability theory. In C. R. Roa and S. Sinharay (Eds.) *Handbook of Statistics: Psychometrics* (Vol. 26, pp. 81-124). Oxford, UK: Elsevier.

White, J. L., & Kratochwill, T. R. (2005). Practice guidelines in school psychology: Issues and directions for evidence-based interventions in practice and training. *Journal of School Psychology, 43,* 99-115.

Williamson, P. & McLeskey, J. (2011). An investigation into the nature of inclusion problem-solving teams. *The Teacher Educator, 46,* 316-334.

Worthington, R. L. & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist, 34*, 806-838.

Zins, J. E., Bloodworth, M. R., Weissberg, R. P., & Walberg, H. J. (2004). The scientific base linking social and emotional learning to school success. In J. E. Zins, R. P. Weissberg, M. C. Wang, & H. J. Walberg (Eds.), *Building academic success on social and emotional learning: What does the research say?* (pp. 3–22). New York: Teachers College Press.

Performance Diagnostic Checklist

Answer each of the following questions, providing data in support of your answer if possible.

Antecedents and Information

Yes	No	
O	O	Is there a written job description telling exactly what is expected of the employee?
O	O	Has the employee received adequate instruction about what to do? (not training - explicit instructions like "I want you to do this, this, and this before we leave today...")
O	O	Are employees aware of the mission of the department/organization? Can they tell you what it is?
O	O	Are there job or task aids in the employees' immediate environment? Visible while completing the task in question? Reminders to prompt the task at the correct time/duration?
O	O	Is the supervisor present during task completion?
O	O	Are there frequently updated, challenging, and attainable goals set that employees are comfortable with/feel are fair?

Equipment and Processes

Yes	No	
O	O	If equipment is required, is it reliable? In good working order? Ergonomically correct?
O	O	Is the equipment & environment optimally arranged in a physical sense?
O	O	Are larger processes suffering from certain incomplete tasks along the way (process disconnects)?
O	O	Are these processes arranged in a logical manner, without unnecessary repetition? Are they maximally efficient?
O	O	Are there any other obstacles that are keeping the employee from completing the task?

Knowledge and Skills

Yes	No	
O	O	Can the employee tell you he/she is supposed to be doing and how to do it? Have they mastered the task? If fluency is necessary, are they fluent?
O	O	Can the employee physically demonstrate completion of the task? Have they mastered the task? If fluency is necessary, are they fluent?
O	O	Does the employee have the capacity to learn how to complete the job?

Consequences

Yes	No	
O	O	Are there consequences delivered contingent on the task? -frequency? (list)_____ -immediacy? (list)_____ -consistency/probability? (list)_____ -positive or negative? (circle one) -Are there premack reinforcers?
O	O	Do employees see the effects of performance? (How? Natural /arranged)
O	O	Do supervisors deliver feedback? (How? Written / verbal; direct /indirect)
O	O	Is there performance monitoring? (Self / supervisor direct / supervisor indirect)
O	O	Is there a response effort associated with performing?
O	O	Are there other behaviors competing with the desired performance?

Figure 2. A sample of a performance diagnostic checklist, currently used by the author in teaching Performance Management at Western Michigan University.

Appendix B. Performance Diagnostic Checklist – Human Services (PDC-HS)

PDC-HS

Employee's Name: _____ Interviewer: _____ Date: _____

Describe Performance Concern: _____

Instructions: Answer the questions below about the employee's specific performance problem (not the employee in general). Items with an asterisk (*) should be answered only after the information is verified through direct observation.

Performance Diagnostic Checklist – Human Services

TRAINING

1.	❍ Yes ❍ No	Has the employee received formal training on this task? If yes, check all applicable training methods: ❍ Instructions ❍ Demonstration ❍ Rehearsal
2.	❍ Yes ❍ No	Can the employee accurately describe the target task and when it should be performed?*
3.	❍ Yes ❍ No	Is there evidence that the employee has accurately completed the task in the past?
4.	❍ Yes ❍ No ❍ N/A	If the task needs to be completed quickly, can the employee perform it at the appropriate speed?*

TASK CLARIFICATION & PROMPTING

1.	❍ Yes ❍ No	Has the employee been informed that he/she is expected to perform the task?
2.	❍ Yes ❍ No	Can the employee state the purpose of the task?*
3.	❍ Yes ❍ No	Is a job aid (ex. a checklist, data sheet) for completing the task visibly located in the task area?*
4.	❍ Yes ❍ No	Is the employee ever verbally, textually, or electronically reminded to complete the task?
5.	❍ Yes ❍ No	Is the task being performed in an environment well-suited for task completion (ex. not noisy or crowded)?

RESOURCES, MATERIALS, & PROCESSES

1.	❍ Yes ❍ No	Are there sufficient numbers of trained staff available in the program?
2.	❍ Yes ❍ No ❍ N/A	If materials (ex., teaching stimuli, preferred items) are required for task completion, are they readily available (ex., easy to find, nearby)? If no materials are required, proceed to question 5.* List materials below and indicate their availability. Item 1: _____ Item 2: _____ Item 3: _____ Item 4: _____
3.	❍ Yes ❍ No	Are the materials necessary to complete the task well-designed?*

	❍ N/A	
4.	❍ Yes ❍ No ❍ N/A	Are the materials necessary to complete the task well organized?*
5.	❍ Yes ❍ No	Is performance suffering from other tasks not being completed first? If so, indicate those tasks below. Task 1: _____ Task 2: _____ Task 3: _____ Task 4: _____
6.	❍ Yes ❍ No ❍ N/A	If you answered YES for Question 5, are other employees responsible for completing any of the earlier tasks in the process? If so, indicate the employee(s) below. Task 1: _____ Task 2: _____ Task 3: _____ Task 4: _____

PERFORMANCE CONSEQUENCES, EFFORT, & COMPETETION

1.	❍ Yes ❍ No	Is the employee ever directly monitored by a supervisor? If so, indicate the frequency of monitoring. ❍ hourly ❍ daily ❍ weekly ❍ monthly ❍ Other: _____
2.	❍ Yes ❍ No	Does the employee ever receive feedback about the performance? If yes, indicate below. By whom? _____ How often? _____ Delay from task? _____ Check all that apply: Feedback Focus: ❍ Positive ❍ Corrective Feedback Type: ❍ Written ❍ Verbal ❍ Graphed ❍ Other: _____
3.	❍ Yes ❍ No	Does the employee ever see the effects of accurate task completion? If yes, how? _____
4.	❍ Yes ❍ No	Is the task particularly effortful or difficult?
5.	❍ Yes ❍ No	Do other tasks appear to take precedence over the target task? If yes, indicate these tasks below. Task 1: _____ Task 2: _____ Task 3: _____ Task 4: _____

Appendix C. *Performance Diagnostic Checklist for Schools* (PDC-S)

Performance Diagnostic Checklist for Schools (PDC-S)

Directions: Please complete this checklist based on your experience with the school-based problem-solving team you are currently working with at this school. For each statement there are ***3 responses required***. The three questions below will be asked each time, please **select only one response** for each of the three items:

1. Is the statement true for your team?
    - Y = Yes
    - N = No
2. If you respond "No" to question #1, please indicate why you think it is not currently true:
    - P = Other needs take priority over this activity/action during team meetings
    - K = Limited knowledge/experience with this activity/action
    - C = Lack of communication/direction provided to team members regarding this activity/action
    - S = Limited support/acknowledgement from administration/team for this activity/action
    - T = Unable to devote time to support this activity/action
3. Regardless of your previous responses, indicate how important you believe the adoption of the practice to be for your teams' overall functioning:
    - V = Very important
    - S = Somewhat important
    - N = Not important

1. In Place?	**2. If "No", Why Not?**	***Area #1 – Training/Knowledge***	**3. How Important?**
Y N	**P K C S T**	1. Team members have been trained on how to establish goals for behavior change.	**V S N**
Y N	**P K C S T**	2. Team members have been trained in the process of analyzing students' behavior problems.	**V S N**
Y N	**P K C S T**	3. Team members received training in data-driven assessment of student behavior problems.	**V S N**
Y N	**P K C S T**	4. One or more team members are trained in the collection and analysis of data that are used in decision-making.	**V S N**
Y N	**P K C S T**	5. Team members have been trained to identify and select behavioral interventions that are evidence-based.	**V S N**
Y N	**P K C S T**	6. Training is available to assist team members and appropriate staff to accurately implement the selected intervention.	**V S N**
Y N	**P K C S T**	7. Team members are able to describe "fidelity of implementation checks" and when they should be performed.	**V S N**
Y N	**P K C S T**	8. The team includes at least one member trained in functional assessment.	**V S N**
1. In Place?	**2. If "No", Why Not?**	***Area #2 – Task Clarification & Prompting***	**3. How Important?**
Y N	**P K C S T**	9. Each core team member has been assigned a specific role within the team (e.g., time keeper, note taker, data analyst, etc.).	**V S N**
Y N	**P K C S T**	10. Team members understand their role on the team. If yes briefly explain your role:	**V S N**
Y N	**P K C S T**	11. Team meetings are viewed as important and as a priority for team members.	**V S N**

Y N	**P K C S T**	12. Team members are verbally or electronically (i.e., email or text) reminded to collect or bring data to share at team meetings.	**V S N**
Y N	**P K C S T**	13. The meeting agenda is communicated in advance of the meeting.	**V S N**
Y N	**P K C S T**	14. The process for conducting team meetings has been clearly established and followed.	**V S N**
Y N	**P K C S T**	15. All team members participate fully during team meetings.	**V S N**
Y N	**P K C S T**	16. All team members contribute their ideas/concerns during team meetings and are heard.	**V S N**
Y N	**P K C S T**	17. Team members develop operational definitions for students' behavioral problems.	**V S N**
Y N	**P K C S T**	18. There is an expectation that team members will bring data on individual students to discuss during the team meetings.	**V S N**
Y N	**P K C S T**	19. By the end of each team meeting, the team ensures that all members understand the selected intervention(s) and their assigned roles.	**V S N**
Y N	**P K C S T**	20. A task analysis of intervention implementation (e.g., a checklist containing all intervention components, data sheet with step-by-step instructions) is provided to the team before implementing the intervention.	**V S N**
Y N	**P K C S T**	21. Team members are encouraged to ask for assistance or clarification prior to beginning intervention implementation as needed.	**V S N**
Y N	**P K C S T**	22. Team members assigned to implement the intervention are verbally or electronically (i.e., email or text) reminded to complete the implementation of the intervention.	**V S N**
Y N	**P K C S T**	23. Team members are verbally or electronically (i.e., email or text) reminded to complete fidelity of implementation checks.	**V S N**
1. In Place?	**2. If "No", Why Not?**	***Area #3 – Resources, Materials & Processes***	**3. How Important?**
Y N	**P K C S T**	24. Team members consistently create behavioral goals for referred students.	**V S N**
Y N	**P K C S T**	25. Team members use collected data to identify the problem.	**V S N**
Y N	**P K C S T**	26. Team members consistently link assessment findings to intervention selection.	**V S N**
Y N	**P K C S T**	27. Team members use collected data to select interventions.	**V S N**
Y N	**P K C S T**	28. Team members consistently select evidence-based interventions to address student problem behavior.	**V S N**
Y N	**P K C S T**	29. Students' progress towards goals is reported at subsequent team meetings.	**V S N**
Y N	**P K C S T**	30. Team members evaluate student progress towards goals using data.	**V S N**
Y N	**P K C S T**	31. Team members document that interventions were implemented according to the plan.	**V S N**
Y N	**P K C S T**	32. Fidelity of intervention implementation checks are completed regularly.	**V S N**
Y N	**P K C S T**	33. Team members use collected data to monitor intervention effectiveness.	**V S N**

Y N	**P K C S T**	34. It is clear to the team when intervention changes are needed.	**V S N**
Y N	**P K C S T**	35. The appropriate materials for intervention implementation are available for staff (e.g., manuals, timers, forms, charts, etc.).	**V S N**
Y N	**P K C S T**	36. Required materials for assessment and intervention implementation are organized (e.g., materials are kept in a specific location, there is a process for checking materials out, etc.).	**V S N**
Y N	**P K C S T**	37. Required materials for assessment and intervention are available.	**V S N**
Y N	**P K C S T**	38. Funding, materials and space are adequate to implement intervention plans.	**V S N**
Y N	**P K C S T**	39. Team meeting notes are provided to team members after the meeting.	**V S N**
Y N	**P K C S T**	40. Team members have the time available to do the work required for their role on the team.	**V S N**
		41. On what basis are interventions most often generated? (Rank order all that apply with 1=most often generated): __Evidence-base __Team member experience __Workshop Idea __Data __Other:_____	**V S N**
1. In Place?	**2. If "No", Why Not?**	***Area #4 – Performance Feedback/Consequences & Outcomes***	**3. How Important?**
Y N	**P K C S T**	42. Team members are recognized and praised when data are used to determine interventions.	**V S N**
Y N	**P K C S T**	43. There are negative consequences/resistance from teachers and/or team members if data are not used to guide intervention decisions.	**V S N**
Y N	**P K C S T**	44. Negative consequences referred to question #42 directly impact team member actions.	**V S N**
Y N	**P K C S T**	45. The use of data to determine intervention decisions is supported and encouraged by all team members.	**V S N**
Y N	**P K C S T**	46. Staff implementing interventions are directly monitored by a team member and provided with support and/or feedback when needed.	**V S N**
Y N	**P K C S T**	47. Staff members responsible for implementing interventions receive feedback from team members about intervention implementation.	**V S N**
Y N	**P K C S T**	48. Team members receive feedback about the implementation or lack of implementation of the intervention from administration.	**V S N**
Y N	**P K C S T**	49. Students referred to the team typically make progress toward the behavioral goal.	**V S N**
Y N	**P K C S T**	50. The team process at this school is effective at helping referred students.	**V S N**
Y N	**P K C S T**	51. The intervention plan developed for referred students is effective at preventing future occurrence of problem behaviors.	**V S N**
Y N	**P K C S T**	52. The team at this school functions well.	**V S N**

Appendix D. Expert Reviewer Demographic Form

Expert Reviewer Demographic Questionnaire

Which of the following best describes your current occupation?
❑ Researcher
❑ Practitioner

How long have you taught, researched, and/or practiced in the field(s) of education and/or mental health?
❑ 0-5 years
❑ 6-10 years
❑ 11-15 years
❑ 16-20 years
❑ 20+ years

How many years of formal training (i.e., didactic courses, supervised practicum experiences) have you received in Applied Behavior Analysis (ABA) including methods of Functional Behavioral Assessment/Analysis (FBA)?
❑ Less than 1 year
❑ 1-2 years
❑ 3-4 years
❑ 5+ years

How many years of experience (i.e., practicing in an applied setting, teaching) have you acquired utilizing principles of ABA including methods of FBA?
❑ Less than 1 year
❑ 1-3 years
❑ 4-7 years
❑ 8-12 years
❑ 13+ years

What is your gender? ❑Male ❑Female

What is your age?
❑ Less than 35
❑ 35-44
❑ 45-54
❑ 55+

Which best describes your ethnicity? ***Check all that apply.***
❑ African-American or Black
❑ American Indian or Alaska Native
❑ Asian American
❑ Caucasian or White
❑ Hispanic or Latino/a
❑ Native Hawaiian or other Pacific Islander
❑ Other: _____

What is the highest academic degree you hold?
❑ High school diploma
❑ Bachelor's
❑ Other: _____
❑ Master's/Specialist
❑ Doctorate

Appendix G. Problem-Solving Team Information Form

Appendix E. Expert Review Form

Dear Expert Reviewer,

This measure is designed to evaluate the content validity of the *Performance Diagnostic Checklist – Schools* (PDC-S). In order to do that, please rate each item as follows.

- Please indicate to which factor the item belongs. The factors are listed along with a definition of each. If you do not think the item belongs with any factor specified, please circle 5.
- Please rate the item's level of representativeness for the assigned factor on a scale of 1-4:
  - 1 = not representative
  - 2 = needs major revisions to be representative
  - 3 = needs minor revisions to be representative
  - 4 = representative
- Please indicate the level of clarity for each item, also on a four-point scale:
  - 1 = not clear
  - 2 = needs major revisions to be clear
  - 3 = needs minor revisions to be clear
  - 4 = clear
- Finally, evaluate the comprehensiveness of the entire measure by indicating if any items should be deleted or added.

Table 1. Factors and Definitions

#	Factor	Definition
1	Antecedents & Information	
2	Equipment & Processes	
3	Knowledge & Skills – Training	
4	Consequences	
5	Other	

Table 2. Reviewer Rankings

#	Item	Factor					Representativeness				Clarity			
1		1	2	3	4	5	1	2	3	4	1	2	3	4
2		1	2	3	4	5	1	2	3	4	1	2	3	4

Comments:

Appendix F. Problem-Solving Team Information Sheet

Problem-Solving Team Information Sheet

For how long has this team been meeting?

❑ 0-3 months
❑ 3-6 months
❑ 6-9 months
❑ 1-2 academic years
❑ 3-4 academic years
❑ 5+ academic years

How frequently does this team meet?

___ time(s) per: ❑Year ❑Semester ❑Quarter ❑Month ❑Week ❑Day

How many permanent members of this team are there? ___

Please list all permanent members' positions in the school (e.g., Principal):

_____
_____
_____
_____
_____
_____

What grade level(s) does this team serve? (Check all that apply.)

❑ Kindergarten
❑ 1st Grade
❑ 2nd Grade
❑ 3rd Grade
❑ 4th Grade
❑ 5th Grade

Appendix G. Problem-Solving Team Member Demographic Questionnaire

Problem-Solving Team Member Demographic Questionnaire

What is your main role in this school? *Check all that apply.*

- ❑ General education classroom teacher
- ❑ Special education classroom teacher
- ❑ Occupational Therapist
- ❑ Physical Therapist
- ❑ Speech and Language Pathologist
- ❑ Reading Specialist
- ❑ Case manager
- ❑ School Psychologist
- ❑ School Guidance Counselor
- ❑ Principal
- ❑ Assistant Principal
- ❑ Other: _____

What grade level(s) do you currently teach or are assigned to in the district? *Check all that apply.*

- ❑ Pre-Kindergarten (Pre-K or 4K)
- ❑ Kindergarten
- ❑ First grade
- ❑ Second grade
- ❑ Third grade
- ❑ Fourth grade
- ❑ Fifth grade

How long have you served on this problem-solving team?

- ❑ Less than 3 months
- ❑ 3-6 months
- ❑ 6-12 months
- ❑ 1-2 years
- ❑ 2-3 years
- ❑ 3+ years

For how many years have you been a school professional, <u>including</u> the present academic year?

_____

For how many years have you been a school professional, <u>including</u> the present academic year at your current school? _____

What is your gender? ❑Male ❑Female

What is your age?

- ❑ Less than 35
- ❑ 35-44
- ❑ 45-54
- ❑ 55+

Which best describes your ethnicity? *Check all that apply.*

- ❑ African-American or Black
- ❑ American Indian or Alaska Native
- ❑ Asian American
- ❑ Caucasian or White
- ❑ Hispanic or Latino/a
- ❑ Native Hawaiian or other Pacific Islander
- ❑ Other: _____

What type of Wisconsin teaching/other license do you hold?

- ❑ Regular teaching license
- ❑ Substitute teaching license
- ❑ School Psychology license
- ❑ Emergency license
- ❑ Administrator Certificate
- ❑ Other: _____

What is the highest academic degree you hold?

- ❑ High school diploma
- ❑ Master's/Specialist

❑ Bachelor's ❑ Doctorate
❑ Other: _____

How many years of formal training (i.e., didactic courses, supervised practicum experiences) have you received in Applied Behavior Analysis (ABA) including methods of Functional Behavioral Assessment/Analysis (FBA)?

❑ Less than 1 year ❑ 3-4 years
❑ 1-2 years ❑ 5+ years

How many years of experience (i.e., practicing in an applied setting, teaching) have you acquired utilizing principles of ABA including methods of FBA?

❑ Less than 1 year ❑ 8-12 years
❑ 1-3 years ❑ 13+ years
❑ 4-7 years

Appendix H. UW-Madison IRB Approval

THE UNIVERSITY of WISCONSIN MADISON

Education and Social/Behavioral Science IRB

4/15/2015

Submission ID number:	2013-0417-CR002
Title:	Functional Behavioral Assessment of Problem-Solving Teams: Adapting the Performance Diagnostic Checklist for Application in Schools
Principal Investigator:	JENNIFER M ASMUS
Point-of-contact:	JENNIFER M ASMUS, THOMAS R KRATOCHWILL, STEPHANIE K SORENSEN
IRB Staff Reviewer:	KAMIE LECLAIR

A designated ED/SBS IRB member conducted an expedited review of the above-referenced continuing review progress report form. The study was approved by the IRB member for the period of 12 months with the expiration date of 4/14/2016. The study qualified for expedited review pursuant to 45 CFR 46.110 and, if applicable, 21 CFR 56.110 and 38 CFR 16.110:

Category 7: Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, or quality assurance methodologies

To access the materials approved by the IRB, including any stamped consent forms and recruitment materials, please log in to your ARROW account and view the documents tab in the submission's workspace.

Please review the Investigator Responsibilities guidance (http://go.wisc.edu/m0lovn) , which includes a description of IRB requirements for submitting continuing review progress reports, changes of protocol and reportable events.

Please contact the appropriate IRB office with general questions: Health Sciences IRBs at 608-263-2362 or Education and Social/Behavioral Science IRB at 608-263-2320. For questions related to this submission, contact the assigned staff reviewer.

Appendix I. Consent Form for Expert Reviewers

UNIVERSITY OF WISCONSIN – MADISON
Department of Educational Psychology
School Psychology Program
1025 West Johnson Street
Madison, WI 53706-1796

Functional Behavioral Assessment of Problem-Solving Teams: Adapting the Performance Diagnostic Checklist for Application in Schools

Dear Sir/Madame:

My name is Stephanie Sorensen, and I am a dissertator in the School Psychology Program at the University of Wisconsin – Madison. I am also a Project Assistant (PA) on the newly funded Institute of Education Sciences (IES) grant entitled, *Systems-Level Analysis of Evidence-Based Intervention Implementation by Problem-Solving Teams,* of which Drs. Thomas Kratochwill and Jennifer Asmus are the Principal Investigators. The goal of this project is to apply principles of functional behavioral assessment/analysis (FBA) to school problem-solving teams (PSTs) to examine barriers associated with teams' selection and implementation of evidence-based interventions (EBIs).

As both part of my role as a PA on the project and that of a dissertator, I am adapting the *Performance Diagnostic Checklist* (PDC) from its original form intended for application in Organizational Business Management and Human Resources, to a measure suitable for application with school-based problem-solving teams (PSTs). The goal of the adapted PDC will be to identify barriers to PSTs' selection and implementation of evidence-based interventions (EBIs), and ultimately use the results to assist in the development and implementation of individualized, function-based interventions for the PSTs to improve their uptake of EBIs. We are writing to invite you to serve as an expert reviewer of our revised PDC. We will use your responses and feedback to revise our measure before piloting them with local PSTs.

If you choose to participate as an expert reviewer, your participation will involve the completion of a demographic questionnaire and one questionnaire to sort items into measures of antecedents and information; equipment and processes; knowledge and skills – training; and consequences. The questionnaires will take about one hour to complete.

Participation in this study is completely voluntary. Should you agree to participate, you are free to stop participating at any time, without penalty. You are also free to choose to not answer any questions with which you do not feel comfortable, and you will not be penalized for any such choice.

All of your responses on the questionnaires will be kept confidential. The questionnaires will only be available to Dr. Kratochwill, Dr. Asmus, and me. We value and want to recognize your time and contribution to the development of our tool; space will be included in my dissertation

and any resulting publications to name those who would like to be recognized by name for their contributions as expert reviewers. There is space on the signature page to consent or decline to be recognized by name as an expert reviewer. In addition, your name will be entered in a drawing for a chance to win one of four $50 gift cards.

There are no anticipated risks associated with your participation as an expert reviewer. There are also no direct benefits for participation. You will be contributing to the development of a new descriptive assessment tool, which may promote PSTs' selection and implementation of EBIs to improve student behavior. You will also be assisting in the completion of this dissertation project.

If you agree to participate in this study, click here ["I AGREE hyperlink]" to begin.

Thank you for considering participating in the dissertation project. Please feel free to contact Stephanie with any questions or concerns by phone (608) 262-3027 or by email at sksorensen@wisc.edu. If you have any questions regarding your rights as a research participant, please contact the Education Research IRB at (608) 263-2320.

Sincerely,

Stephanie K. Sorensen, M.S.
Dissertator, School Psychology
335A Educational Sciences
1025 West Johnson Street
Madison, WI 53706
sksorensen@wisc.edu
(608) 262-3027

Thomas R. Kratochwill, Ph.D.
Professor, School Psychology
333 Educational Sciences
1025 West Johnson Street
Madison, WI 53706
tomkat@education.wisc.edu
(608) 262-5912

Jennifer Asmus, Ph.D.
Professor, School Psychology
316B Educational Sciences
1025 West Johnson Street
Madison, WI 53706
asmus@wisc.edu
(608) 262-3027

Appendix J. Consent Form for Problem-Solving Team Members

UNIVERSITY OF WISCONSIN – MADISON
Department of Educational Psychology
School Psychology Program
1025 West Johnson Street
Madison, WI 53706-1796

Functional Behavioral Assessment of Problem-Solving Teams: Adapting the Performance Diagnostic Checklist for Application in Schools

Dear Problem-Solving Team Member:

My name is Stephanie Sorensen, and I am a dissertator in the School Psychology Program at the University of Wisconsin – Madison. I am also a Project Assistant (PA) on the newly funded Institute of Education Sciences (IES) grant entitled, *Systems-Level Analysis of Evidence-Based Intervention Implementation by Problem-Solving Teams*, of which Drs. Thomas Kratochwill and Jennifer Asmus are the Principal Investigators. The goal of this project is to apply principles of functional behavioral assessment/analysis (FBA) to school problem-solving teams (PSTs) to examine barriers associated with teams' selection and implementation of evidence-based interventions (EBIs).

As both part of my role as a PA on the project and that of a dissertator, I am adapting the *Performance Diagnostic Checklist* (PDC) from its original form intended for application in Organizational Business Management and Human Resources, to a measure suitable for application with school-based problem-solving teams (PSTs). The goal of the adapted PDC will be to identify barriers to PSTs' selection and implementation of evidence-based interventions (EBIs), and ultimately use the results to assist in the development and implementation of individualized, function-based interventions for the PSTs to improve their uptake of EBIs. We are writing to invite you to serve as a participant in the piloting of the PDC.

If you choose to participate, you will be asked to complete a demographic questionnaire and the PDC at two separate times. The completion of these measures will take about 45 minutes over two sessions for a total time of approximately 1.5 hours. Participation in this study is completely voluntary. Should you agree to participate, you are free to stop participating at any time, without penalty. You are also free to choose to not answer any questions with which you do not feel comfortable, and you will not be penalized for any such choice.

There are no anticipated risks associated with your participating in this study. All of your responses on the questionnaires will be kept confidential. A code will be automatically entered on all completed forms so that no identifiable information will be linked to any of your responses. We value and want to recognize your time and contribution to the development of our tool, so we will be entering all participants' names into a drawing from which we will select 8 PST members to earn a $50 gift card.

While there are no direct benefits for participation in this study, you will be contributing to the development of a new descriptive assessment tool, which may promote PSTs' selection and implementation of EBIs to improve student behavior. You will also be assisting in the completion of this dissertation project.

If you agree to participate in this study, click here ["I AGREE hyperlink]" to begin.

Thank you for considering participating in the dissertation project. Please feel free to contact Stephanie with any questions or concerns by phone (608) 262-3027 or by email at sksorensen@wisc.edu. If you have any questions regarding your rights as a research participant, please contact the Education Research IRB at (608) 263-2320.

Sincerely,

Stephanie K. Sorensen, M.S.
Dissertator, School Psychology
335A Educational Sciences
1025 West Johnson Street
Madison, WI 53706
sksorensen@wisc.edu
(608) 262-3027

Thomas R. Kratochwill, Ph.D.
Professor, School Psychology
333 Educational Sciences
1025 West Johnson Street
Madison, WI 53706
tomkat@education.wisc.edu
(608) 262-5912

Jennifer Asmus, Ph.D.
Professor, School Psychology
316B Educational Sciences
1025 West Johnson Street
Madison, WI 53706
asmus@wisc.edu
(608) 262-3027