Alexander B. Brewer

Education

Indiana State University
Ph.D. Student in School Psychology
Academic Advisor: Leah Nellis, Ph.D.
August 2010

Indiana State University
Master of Education
August 2005

Indiana University – Purdue University at Indianapolis
B.S. Elementary Education with endorsement in Social Studies grades 1-9, December 2003

Current Employment

August 2009 to Present **Metropolitan School District of Martinsville, Martinsville, Indiana**

Advanced School Based Internship

August 2008 to May 2009 **Metropolitan School District of Decatur Township – Indianapolis, Indiana**

APA Accredited Pre-Doctoral Internship

May 2007 to August 2008 **Houston Independent School District – Houston, Texas**

Research Experience

July 2010 **A Descriptive Study of a Building-Based Team Problem-Solving Process**
Indiana State University
Terre Haute, Indiana

- A dissertation Presented to The College of Graduate and Professional Studies Department of Communication Disorders and Counseling, School, and Educational Psychology
- Advisor: Dr. Eric Hampton, Indiana State University

Professional Organizations/ Certifications

- National Association of School Psychologists (student member)
- APA division 16 student member
- Indiana Association of School Psychologists (student member)
- National Organization of Victims Assistance: National Crisis Responder Certification
- Crisis Prevention Institute: Nonviolent Crisis Intervention Certification

A DESCRIPTIVE STUDY OF A BUILDING-BASED TEAM PROBLEM-SOLVING PROCESS

A dissertation

Presented to

The College of Graduate and Professional Studies

Department of Communication Disorders

and Counseling, School, and Educational Psychology

Indiana State University

Terre Haute, Indiana

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

Alexander B. Brewer

August 2010

Keywords: Education, Intervention, Problem-Solving Team

COMMITTEE MEMBERS

Committee Chair: Eric Hampton, Ph.D.

Associate Professor of Department of Communication Disorders and Counseling, School, and Educational Psychology

Indiana State University

Committee Member: Leah Nellis, Ph.D.

Associate Professor of Department of Communication Disorders and Counseling, School, and Educational Psychology

Indiana State University

Committee Member: Bridget Roberts-Pittman, Ph.D.

Assistant Professor of Department of Communication Disorders and Counseling, School, and Educational Psychology

Indiana State University

ABSTRACT

The purpose of this study was to empirically evaluate Building-Based Teams for General Education Intervention or BBT for GEI. BBT for GEI is a team problem-solving process designed to assist schools in conducting research-based interventions in the general education setting. Problem-solving teams are part of general education and provide support to students with academic or behavioral concerns by creating individualized interventions that teachers can use in the classroom. Historically, problem-solving teams' two primary goals were to reduce referrals to special education and improve student performance on academic or behavioral concerns.

This study examined the effectiveness of BBT for GEI by analyzing BBT for GEI teams' alignment with the best practice indicators of intervention design and by evaluating how BBT for GEI teams' practices predict student outcome. The analysis was done by reviewing permanent products of team GEI practices submitted by elementary school problem-solving teams trained in the BBT for GEI process by the Blumberg Center for Interdisciplinary Studies in Special Education. The teams' permanent products were rated on 13 quality indicators of intervention design using a Likert type scale of 1-5 on adherence and presence of the indicator. The higher the rating on the scale, the greater the alignment with the identified best practices for that indicator. The quality indicators include the following: (a) behavioral definition, (b) baseline data, (c) problem validation, (d) problem analysis, (e) goal setting, (f) delivery specifics, (g) empirically-supported content variables, (h) measurement strategy, (i) decision-making plan, (j) progress monitoring, (k) formative evaluation, (l) treatment integrity, and (m) summative evaluation.

The average indicator ratings ranged from a low of 1.44 to a high of 3.64. This range suggests that the teams implemented some of the best practice indicators to a high degree, while other indicators were either not implemented to a high degree or not addressed. BBT for GEI teams implemented the Problem Analysis and Plan Development components with the highest fidelity while implementing the Plan Implementation and Plan Evaluation components with the lowest fidelity. When analyzing the themes and commonalities, it became apparent that many teams did not conduct more than their initial meeting in order to implement and monitor a plan. In addition to the 13 indicator ratings, two student outcome ratings were also assigned to teams' permanent products, Goal Attainment Scaling (GAS) and Student Measured Performance (SMP). The average rating for GAS was 2.92. The average for SMP was 1.93. Two multiple regression analyses were conducted to determine the effect the 13 quality indicators have on GAS and SMP. The linear combination of the quality indicators of intervention design ratings was significantly related to both GAS and SMP. Individually, Intervention Plan Development and Problem Analysis were significant predictors of GAS. Four indicators were significant predictors of SMP, Problem Validation, Goal Setting, Intervention Plan Development, and Formative Evaluation.

ACKNOWLEDGMENTS

I consider myself very fortunate to have the support and love of many friends and family whose support, love, and encouragement I could not have succeeded without. First and foremost, I would like to thank God for the many blessings he has given me. To my mom and siblings, Jane, Natasha, Ty, and Nick, you have provided me the foundation from which I have been allowed to grow both personally and professionally. For this I thank you. To my wife Jennifer, I am not able to find the words to express how much you have meant to me. Your love and support has been immeasurable. I'd also like to mention my father Paul, who I know would be proud of me, and my son Adam, who has been my inspiration.

This project would not have been completed without the help and guidance of my committee members, Drs. Eric Hampton, Leah Nellis, and Bridget Roberts-Pittman. I would also like to also thank the staff of the Blumberg Center for opening their office doors and filing cabinets to let me have access to the data that made this study possible. I would like to send a final thank you to Ke'Shana and Megan for their help in conducting the reliability checks for the data analysis.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Problem-solving teams, or PSTs, have functioned in the schools under many different names in the past few decades, and the function of each is essentially the same. Lane, Pierson, Robertson, and Little (2004) described the problem-solving intervention process these teams conduct as a function of general education that provides support to students with academic or behavioral concerns while maintaining the student's least restrictive environment. Fuchs, Fuchs, and Bahr (1990) described problem-solving intervention as a teacher's modification to instruction or the environment to help manage a "difficult-to-teach" student in a general education setting. Problem-solving intervention team processes in the schools are consistent with where interventions are carried out (general education classroom) and the types of problems targeted through intervention (academic and behavioral concerns).

One particular method for conducting problem-solving intervention teams is termed Creative Problem Solving for General Education Intervention (CPS for GEI; Bahr et al., 2006). CPS for GEI is a process based on the work of Osborn and Parnes in the field of creativity (Bahr et al., 2006). The term general education intervention is used commonly in the state of Indiana and is the primary activity carried out by problem-solving teams. According to the Indiana State Board of Education's Special Education Rules, general education intervention is not a prerequisite for referring a student for an evaluation (Indiana State Board of Education, 2002).

Rather, Indiana's Special Education Rules define general education intervention as "a written formal system, at the building level, of methods and procedures used with a student to address those aspects of a student's classroom performance that are substantially affecting general education outcomes" (Indiana State Board of Education, 2002, p. 8). The goal of CPS for GEI is to provide teams with a structured method of facilitating groups for the purpose of making them more effective and more efficient. In a study by Hampton (2004), 11 focus groups were conducted with six school based problem-solving teams from elementary school teams and five school district trainers of CPS for GEI teams. Of these, respondents reported four primary strengths of the CPS for GEI process. The respondents reported that the CPS for GEI process was more effective than other processes they had used in the past. They also noted that the process identified a child's strengths. An additional benefit was the process's ability to impact all of the children in a school, not just the identified child. Lastly, the focus groups revealed that the process led to focus by its group members.

The CPS for GEI process has recently been revised to align more closely with the behavioral consultation model (Bergan & Kratochwill, 1990) and was renamed Building-Based Teams: Problem Solving to Support General Education Intervention (BBT for GEI) in 2006. In addition to becoming aligned more closely with the behavior consultation model, additional changes include an increased focus on the following: (a) assessment, (b) progress monitoring data, (c) process support, and (d) flexibility in team membership. As the BBT for GEI process was first implemented beginning in the fall of 2006, no research has been located which evaluates the effectiveness of BBT for GEI teams at this time.

This study examined the effectiveness of BBT for GEI through objective, permanent product indices of intervention quality, an analysis of BBT for GEI's effect on student

performance outcomes, and BBT for GEI's use of quality indices of intervention design. Researchers have attempted to relate the quality indices of intervention design to student outcomes (Flugum & Reschly, 1994; Strickler, 2004; Telzrow, McNamara, & Hollinger, 2000; Upah, 1998). However, according to Upah (1998), few studies have empirically evaluated the impact of the quality indices on student outcomes. This study contributes to this emerging body of research by assessing the relationship between the quality indicators of intervention design and student outcomes of BBT for GEI teams.

Need

The need for specialized assistance within the schools is substantial. Students face a growing number of risk factors associated with school failure. These risks include, among others, poor grades, retention, increased poverty, and frequent moving between schools (Horn & Carroll, 1997). Reading concerns remain the most frequent reason for special education referrals (Snow, Burns, & Griffin, 1998), but Baker, Gersten, and Grossen (2002) argued those reading concerns may be due to a lack of quality instruction and intervention in reading. In an attempt to address these students at risk for failure, educators have sought support to identify instructional methods to accommodate students' specific needs. A survey of 200 elementary schools across the United States found that 85% of that support has come in the form of problem-solving intervention teams (Slonski-Fowler & Truscott, 2004).

Researchers have evaluated student outcomes from interventions designed by problem-solving intervention teams (Bahr et al. 2006; Chalfant & Pysh, 1989; Fuchs, Fuchs, & Bahr, 1990; Gravois & Rosenfield, 2002). Student outcome data used by researchers to assess problem-solving teams' success includes reduction in referrals to special education and improved student performance on academic or behavioral concerns (Sindaler, Griffen, Smith, &

Wantanabe, 1992). A few researchers have examined the relationship between research-identified quality indices of intervention and student outcomes. These researchers have found a linear correlation between an increased use of the quality indices of intervention and an increase in positive student outcomes (Flugum & Reschly, 1994; Strickler, 2004; Telzrow et al., 2000; Upah, 1998). Further research is necessary to support these findings. Because BBT for GEI has yet to be extensively empirically evaluated, there is a need to investigate both the student outcomes and the quality indices of intervention. Additional support for including quality indices will be addressed through evaluating the link between those quality indices of intervention design and student outcomes of BBT for GEI teams.

Purpose Statement

BBT for GEI teams are designed to provide collaborative problem solving for intervention design in the general education setting. The goal of collaborative problem solving within the schools is to design strategies that increase success for students in the classroom (Allen & Graden, 2002). The purpose of the current study was to evaluate empirically the BBT for GEI teams' effectiveness at improving student outcomes. In order to do this, the BBT for GEI teams were assessed on two levels, process and outcome. To assess the BBT for GEI process, the degree to which BBT for GEI teams implement quality indices of intervention design were examined. To assess the outcome of the BBT for GEI teams, the following student outcome variables were examined: (a) rates of referral to special education, (b) student performance on academic or behavioral targeted outcomes, and (c) student goal attainment.

Research Questions

The following research questions guided this study:

1. How well do BBT for GEI teams' practices align with the quality indicators of intervention identified in the literature? What themes or commonalities are evident related to BBT for GEI teams' alignment with the quality indicators?

2. What effects do the BBT for GEI teams have on student outcomes? (a) To what degree do BBT for GEI teams' alignment with the quality indicators of intervention predict the attainment of goals set by BBT for GEI teams? (b) To what degree do BBT for GEI teams' alignment with the quality indicators of intervention predict student performance on academic or behavioral measures implemented during GEI?

3. Do significant differences exist between the quality indicators of the interventions created by BBT for GEI teams and the outcomes for the following student groups: (a) students referred for a special education evaluation and subsequently identified as being eligible for special education, (b) students referred for a special education evaluation and subsequently not found eligible for special education, and (c) students no longer in GEI?

Definitions

BBT for GEI

Building-Based Teams: Problem Solving to Support General Education Intervention is a process modified from CPS for GEI and utilized for conducting a problem-solving intervention team process.

CPS for GEI

Creative Problem Solving for General Education Intervention is a team-based process for conducting the problem-solving process to address student needs.

Innovation Configuration

A tool for research and evaluation that allows researchers to rate skills in observable and measurable terms; the format for rating level of implementation is based on the framework of Hord, Rutherford, Huling-Austin, and Hall (1987).

Quality Indices for Intervention Design

Quality indices for intervention design are elaborations on or components of the behavioral consultation model (Bergan & Kratochwill, 1990), which have been identified in the literature. This study included the following quality indices for intervention design from Tilly and Flugum (1995), Upah and Tilly (2002), and Strickler (2004): (a) behavioral definition, (b) baseline data, (c) problem validation, (d) problem analysis, (e) intervention plan development, (f) strength of intervention, (g) measurement strategy, (h) decision-making plan, (i) progress monitoring, (j) formative evaluation, and (k) summative evaluation.

Problem-Solving Intervention Teams

Problem Solving intervention teams are a function of general education that provide support to students with academic or behavioral concerns while maintaining the student's least restrictive environment (Lane et al., 2004).

CHAPTER 2

LITERATURE REVIEW

The enactment of the Education for All Handicapped Children Act (EAHCA, 1975) provided protection of rights and services to children through special education. The Individuals with Disabilities Education Act (IDEA, 1990) and its reauthorization in 1997 (IDEA, 1997) placed focus on improving those services and rights for eligible children. Prevalence rates of special education placement have increased, most notably in the eligibility of learning disabilities. These prevalence rates have increased 283% between 1976-1977 and 1998-1999 (United States Department of Education, 1998). According to Gresham (2001), the three-step process (referral, psychological evaluation, and team decisions) currently used to identify students as learning disabled too often results in false positives, false negatives, inappropriate placements, or some combination due to behavior problems, poor instruction, language difficulties, and many other reasons.

To reduce this increased number of students being identified for special education including specific learning disabilities, schools began forming school-based teams to address concerns before a referral to special education is made. In fact, in 1989 a national survey reported that 23 states required and 11 others recommended some form of problem-solving team consultation prior to a child being referred for a full special education evaluation (Carter & Sugai, 1989). By 2005, 34 states required problem-solving team consultation and another nine

states recommended the use of problem-solving intervention teams (Truscott, Cohen, Sams, Sanborn, & Frank, 2005). Often, educators have considered these teams to serve only as perfunctory last steps before an inevitable referral for an evaluation for special education eligibility (Kovaleski, 2002). The role of problem-solving teams, however, has recently changed. The reauthorization of the Individuals with Disabilities Education Improvement Act (IDEIA, 2004) further supported the use of alternative methods of delivering services to underachieving students by allowing schools the option of identifying students with specific learning disabilities based on their responsiveness to empirically supported interventions. This paradigm shift allowed schools to use a response to intervention processes for identifying students for special education services when students did not respond to the interventions implemented in the problem-solving process. The shift also allowed schools to address student underachievement as soon as it is evident.

Best Practices for Problem-Solving Intervention Teams

The use of teams to conduct problem-solving intervention began in the 1980s as a method for reducing the number of inappropriate referrals to special education. Since then, best practices for conducting problem-solving intervention teams have been identified in the literature.

Team Composition

The first consideration for a problem-solving team is team membership (Kovaleski, 2002). Researchers have debated the usefulness of the team format. Rosenfield and Gravois (1996) observed that the quality of decisions made by problem-solving teams was no better than what the group's best member could have made. Rosenfield and Gravois attributed these lower quality decisions to the teams' over-emphasis on process, team-member training, and systems issues. In contrast, Chalfant and Pysh (1989) argued that a team format allowed for the

development of a sense of mission and team spirit, which is necessary if schools are to facilitate change in general education.

Variance exists between the number and type of professionals who participate in problem-solving intervention teams. Team composition is, therefore, somewhat difficult to assess in a broad, generalizable manner. Only seven states recommend or specify team composition (Truscott et al., 2005). Bahr, Whitten, Dieker, Kocarek, and Manson (1999) compared problem-solving intervention teams from three Midwestern states and reported that 66% of the team members were support staff (e.g., special educator, school psychologist, social worker, etc.), 19% were general education teachers, and administrators comprised the remaining 15%. The size of the teams in this study averaged between five and seven members.

Administrative Support

The involvement of the school principal in problem-solving intervention teams also has been identified as an important feature because of the need for instructional personnel and support services to collaborate. According to Kruger, Struzziero, Watts, and Vacca (1995), teacher satisfaction with the problem-solving intervention process was related to the perceived level of administrative support. Additionally, team members perceived higher levels of administrative support when the principal actually participated in the problem-solving intervention process (Rosenfield & Gravois, 1996). Rosenfield and Gravois proposed that principals should share with their problem-solving intervention teams their vision of the school and how the teams fit within that vision. Principal support is important in facilitating the transition from the traditional test-and-place paradigm to the problem-solving intervention model for intervening with struggling students (Kovaleski, 2002).

Training and Support for Teams

Researchers have found that, in order for intervention procedures to be sustained, problem-solving team members need to be provided with professional development training and support in implementing the process. Vaughn, Hughes, Schumm, and Klinger (1998) conducted a study to examine the effects of professional development on intervention sustainability. Nine teachers were trained in conducting reading and writing interventions for students with learning disabilities (e.g., peer tutoring and letter-sound association). Upon completion of the training, all of the nine teachers had implemented components from each of the interventions. In a one-year follow up, the participating teachers reported on the factors that contributed to the continuing use of the interventions. These factors included: (a) strategies that could be used with the entire class, (b) strategies that improved performance on the standardized state assessment, (c) training that provided the teachers with a global or theoretical understanding of the interventions, and (d) follow-up meetings between the trainers and the teachers.

Ivarie and Russell (1992) studied the impact of 16 hours of training in collaborative consultation for school-based teams on referrals to special education. The training consisted of the following primary objectives: (a) principles of collaborative consultation, (b) prereferral intervention training, and (c) problem-solving process training. Ivarie and Russell found that the teams trained in collaborative consultation increased appropriate special education referrals. Yetter and Doll (2007) investigated the impact logistical resources had on the acceptability of a particular problem-solving process to school staff. Yetter and Doll found that having sufficient staff training in the team procedures increased staff acceptability of the consultation team process.

Bahr et al. (2006) studied the impact of training on problem-solving teams by randomly assigning 24 teams into either a training condition or a control group. The problem-solving teams in the training condition received the following training over the course of one school year: (a) two 1-day intensive process training (at beginning and end of year), (b) a workshop to the entire school staff on problem-solving team process, (c) consultation visits from trainers for team support, and (d) follow-up support following the training year to address individual team needs. The control group received no training. Based on team outcome surveys measuring team effectiveness, meeting timeliness, and intervention practices, Bahr et al. reported that teams receiving the one-year training significantly outperformed untrained teams across outcome variables.

BBT for GEI teams receive similar training as the participant teams in Bahr et al. (2006). Teams utilizing the BBT for GEI process, reviewed later in this chapter, receive an intensive, initial training. The trained personnel then conduct a school wide in-service for the school staff prior to fully implementing the process. The BBT for GEI teams are then provided follow-up training focused on previous implementation outcomes. Technical assistance provided by university trainers is also made available to BBT for GEI teams based on team needs (Blumberg Center for Interdisciplinary Studies in Special Education [Blumberg Center], 2006). The effectiveness of the training for BBT for GEI has not yet been extensively researched.

Timeliness

Due to a limited amount of time for meetings, it is important for teams to utilize procedures to conduct meetings in a timely manner. Team members have cited time constraints as a problem area for problem-solving intervention teams (Lane et al. 2004; Slonski-Fowler & Truscott, 2004). Iverson (2002) made suggestions to improve meeting timeliness: (a) utilize a

timekeeper, (b) prepare for meetings by completing paperwork prior to the meeting, and (c) divide meeting times appropriately. Other strategies for time management include designating roles, posting a meeting agenda with time limits on discussion topics, and establishing a meeting goal (Macan, 1994).

Accountability

Because it is essential that the activities of the problem-solving intervention teams be administered effectively, a system of accountability is needed to ensure that teams are demonstrating effective processes to address teacher referrals (Kovaleski, 2002). Slonski-Fowler and Truscott (2004) reviewed teacher dissatisfaction with the problem-solving intervention process and found that teachers became disengaged from the process when teams demonstrated little accountability for implementation or outcomes. According to Slonski-Fowler and Truscott, the teachers described accountability as a joint responsibility between the referring teacher and the team when the student was not making progress with the implemented intervention.

Collaborative Problem-Solving Process

According to Allen and Graden (2002), problem-solving teams should utilize the four-step behavioral consultation model of problem solving suggested by Bergan (1977) and Bergan and Kratochwill (1990). The four steps of this problem-solving model are: (a) problem identification, (b) problem analysis, (c) plan implementation, and (d) plan evaluation. Fuchs and Fuchs (1989) conducted a study of three teacher groups using this behavioral consultation model. Each group included different combinations of the model steps in order to evaluate the importance of including all four steps of the behavioral consultation model during consultation. The three groups were (a) problem identification and problem analysis; (b) problem identification, problem analysis, and plan implementation; and (c) problem identification,

problem analysis, plan implementation, and plan evaluation. Forty-eight teachers participated in the study and were randomly assigned into one of the three groups or a control group. Pretreatment and posttreatment teacher ratings indicated that reductions in problem behaviors were more pronounced when more components of behavioral consultation were present. This research suggests that, in order for behavioral consultation to be most effective, all four stages should be present. The following sections will discuss in greater detail each of the four stages of behavioral consultation.

Problem Identification

Bergan (1977) defined problem identification as determining the priority of concerns in specific, operational, and behavioral terms. To meet these criteria, the problem must be defined in observable terms that can be directly measured. If multiple problems exist, the team must prioritize problems or identify "keystone" variables or behaviors that, if changed, would impact a larger set of problems (Allen & Graden, 2002).

Problem Analysis

Problem analysis consists of identifying and validating a problem's existence, identifying variables both from the individual and within the classroom that may be of assistance, and developing an intervention plan (Allen & Graden, 2002). To do this, baseline data regarding problem frequency, duration, and intensity must be collected. Also, antecedents of the problem behavior and consequences that may be maintaining the behavior need to be identified (Zins & Erchul, 2002). Zins and Erchul suggested assessing relevant environmental factors as part of problem analysis. Shinn (2002) described Curriculum Based Measurements (CBMs) as one method used to identify specific student academic problems, suggest interventions, and monitor student progress throughout the implementation of the intervention. For example, a team may

use a CBM when conducting a reading intervention to increase a student's reading fluency. First, the team could develop reading probes using passages from the student's reading level. Then, at the beginning of the intervention implementation and throughout implementation, a team member may have the student read a passage and record how many words he or she correctly reads in one minute. Thus, CBMs can be used in problem analysis and in plan implementation. Other assessment techniques that may be used in either problem analysis or plan implementation are single-case experimental designs (Steege, Brown-Chidsey, & Mace, 2002). According to Steege et al., single-case designs are intended to study single units of behavior before, during, and following the introduction of an intervention. Steege et al. identified three reasons why single-case designs are desirable within a problem-solving model. The three reasons are: (a) data can be analyzed easily, (b) objective determination of student progress is documented, and (c) teams can quickly identify the effective and ineffective components of an intervention through ongoing assessment.

Plan Development

Plan development tasks are often considered to be part of the problem analysis phase for problem-solving teams (Allen & Graden, 2002). The primary function of problem-solving teams is to develop and implement general education interventions; therefore, plan development tasks deserve special consideration. Kratochwill, Elliott, and Callan-Stoiber (2002) identified the following three components that should be present during plan development: (a) delivery specifics, (b) goal setting, and (c) empirically supported content variables. The first of these components, delivery specifics, are details including who will be carrying out the intervention, where and when the intervention will be carried out, and materials necessary for implementation (Kratochwill et al., 2002). The second component, goal setting, is the process of defining the

desired level of achievement from intervention implementation (Upah & Tilly, 2002). The third and final component of plan development, empirically supported content variables, signifies the importance of using empirically supported strategies when planning interventions.

Plan Implementation

As previously mentioned, curriculum-based measurement and single-case design interventions can be used within the plan implementation stage. Plan implementation involves the implementation of a plan as intended, continuous monitoring of progress, and change within the plan if necessary (Allen & Graden, 2002). Once an intervention has been identified, the team creates an intervention plan that specifies every aspect of the intervention to be carried out (Allen & Graden, 2002). The best practices for intervention plans are explained in further detail in a subsequent section.

Problem Evaluation

The final step to the problem solving model described by Allen and Graden (2002) is problem evaluation. In this step, the intervention is evaluated and, if it has been ineffective, modifications are designed and implemented. In the problem evaluation stage, teams should use the same monitoring and assessment methods from the problem analysis and plan implementations so that the data is continuous (Zins & Erchul, 2002).

Models for Problem-Solving Teams

In a national survey of problem-solving team practices, Buck, Polloway, Smith-Thomas, and Cook (2003) reported that 47% of states do not use a standard term for problem-solving intervention teams. Despite the many different names given to problem-solving intervention teams, Iverson (2002) reported that the structure of the teams can be divided into one of two models: the case management model and the broad participation model.

The case management model requires a team of guidance counselors, administrators, special educators, school psychologists, and general education teachers. The team receives a referral from a teacher and assigns the case to one of the team members to serve as the case manager. That manager is chosen based on his or her particular fit for the referring case. The case manager then becomes the consultant who brings teachers and parents together for the problem-solving process (Iverson, 2002). The instructional consultation team model (Rosenfield & Gravois, 1996) and the problem-solving intervention (PI) model (Graden, Casey, & Christenson, 1985) are examples of these types of teams.

Broad participation model teams have been described by Iverson (2002) as typically meeting once per week at a regularly scheduled time. These teams cover all of the referral cases at these meetings and process them as a group. Support personnel such as school psychologists and speech language pathologists are not usually involved in these meetings (Iverson, 2002). Mainstream assistance teams, or MATs (Fuchs, Fuchs, & Bahr, 1990), and Teacher Assistance Teams (Chalfant & Pysh, 1989) are examples of these types of teams.

The current study focused on a particular broad participation model referred to as Building-Based Teams: Problems Solving to Support General Education Intervention (BBT for GEI). BBT for GEI represents a recent modification to another broad participation model, Creative Problem Solving for General Education Intervention (CPS for GEI). CPS for GEI, however, was preceded by other broad participation models. According to Burns, Vanderwood, and Ruby (2005), most problem-solving team models are aligned with one of five models, of which two, Teacher Assistance Teams (Chalfant & Pysh, 1989) and Mainstream Assistance Teams (Fuchs, Fuchs, & Bahr, 1990), are broad participation models. Because of this, Teacher

Assistance Teams and Mainstream Assistance Teams are described below to exemplify the processes used in broad participant modeled teams.

Teacher Assistance Teams

Teacher Assistance Teams, or TATs, represent the first generation of teams that were created to provide problem-solving. These teams consisted primarily of teachers and were created to address the needs of struggling students who did not meet the criteria for special education (Chalfant & Pysh, 1989). The goal of TATs is to "provide a forum where classroom teachers can meet and engage in a positive, productive, collaborative, problem-solving process to help students indirectly…through teacher consultation" (Chalfant & Pysh, 1989, p. 50).

The first stage of the TAT process (Chalfant & Pysh, 1989) includes accurately and concisely describing the student's needs and the present classroom problems. In the second stage, the TAT targets student needs through a visualization diagramming process. The third stage of TAT consists of conducting interviews with teachers to collaboratively select and write realistic intervention goals for the referred student. Fourth, teams conduct "efficient and effective" 30-minute meetings for problem-solving. In the fifth stage, the TAT assists the teacher in brainstorming practical strategies for use in his or her classroom. The final stage of TAT consists of developing procedures for measuring intervention effectiveness and providing additional support for the teacher as needed (Chalfant & Pysh, 1989).

The TAT model has many of the best practice components for conducting problem-solving intervention teams. It utilizes the four stages of behavioral consultation described in Allen and Graden (2002), and it utilizes time-management procedures. Limitations identified with the TAT model include inconsistent administrative support for team's decisions and the quality of intervention implementation (Chalfant & Pysh, 1989). An additional limitation is that,

although the TAT model does have a plan implementation stage, it does not maintain best practices within this stage as it does not specify procedures for the intervention plans created by the TAT.

Mainstream Assistance Teams

MATs were developed between 1985 and 1986 in four metropolitan middle schools (Fuchs, Fuchs, & Bahr, 1990). MATs consisted of multidisciplinary school-based teams that were trained in behavioral consultation. MATs were created to bridge special education with general education by including group members from special education and general education as well as other specialists (Fuchs, Fuchs, & Bahr, 1990).

The process used by MATs is an adapted version of behavioral consultation. It borrows its stages from the consultation models of Bergan (1977) and Cantrell and Cantrell (1980) and provides them in scripts for the group to use during consultation (Fuchs, Fuchs, & Bahr, 1990). The process consists of four stages that are implemented in a collaborative manner by the problem-solving intervention team. The first stage is the problem-identification stage and consists of defining the problem in behavioral terms and estimating its frequency, intensity, and duration. In the second stage of problem-analysis, the group determines how the problem behavior is exhibited and why the problem behavior is occurring. The third stage is plan implementation, which consists of monitoring the implementation of the plan and giving corrective feedback. The fourth and final stage in the MAT process includes problem evaluation. In this stage, the teacher and the team meet to assess the effectiveness of the plan and make any necessary modifications (Fuchs, Fuchs, & Bahr, 1990).

Fuchs, Fuchs, Gilman, et al. (1990) reported that the team process was too time consuming and the team model was reduced to one consultant. The one consultant approach,

according to Fuchs, Fuchs, Gilman, et al. (1990) was limited by teacher reports of the interventions being too restrictive and ungeneralizable. For a more in-depth discussion regarding the transition from the use of consultant-teacher teams to the one-on-one behavioral consultation MAT, the reader is referred to Fuchs, Fuchs, Gilman, et al.

Because this model was created to reflect the behavioral consultation model originally described by Bergan (1977), it is consistent with the best practices for the problem-solving intervention process. Researchers have debated whether single consultants are more effective than team consultants at problem solving (Rosenfield & Gravois, 1996; Kovaleski, 2002). Thus, the implications of the MATs' use of single consultants on best practices are currently undetermined.

Building-Based Teams for General Education Intervention

As was previously mentioned, CPS for GEI was renamed Building-Based Teams: Problem Solving to Support General Education Intervention or BBT for GEI (Blumberg Center, 2006). The goal of CPS for GEI was to provide teams with a structured method of facilitating groups for the purpose of making them more effective and more efficient. The CPS for GEI problem-solving meeting consisted of 10 steps: (a) compile background information, (b) open meeting with introductions or role clarifications, (c) review the "current" reality for the student, (d) confirm the desired outcome, (e) generate critical questions, (f) focus on key questions, (g) generate interventions to address key questions, (h) identify promising interventions, (j) develop a plan of action, and (k) wrap-up (Blumberg Center, 2006). BBT for GEI was created in order to reflect team member feedback (Hampton, 2004) and current legislation outlining new methods for identifying eligibility for special education found in the Individuals with Disabilities Education Improvement Act (IDEIA, 2004).

BBT for GEI addresses the limitations of the CPS for GEI by including the problem analysis and plan evaluation stages that were not adequately addressed in CPS for GEI. The added problem analysis stage, termed Review Data and Determine Current Reality for the Student, takes the place of CPS for GEI's third and fourth stages. In this new stage, the team is asked to confirm the area of concern (academic or behavioral) and identify the priority subskills of this concern. The team then identifies any barriers to the student's potential for making progress (e.g., curriculum or school and home environment). After the team has identified potential barriers, the team generates and selects the interventions to be implemented. The team then creates measurable goals for each intervention and subskill selected (Blumberg Center, 2006). The plan implementation stage builds on this stage by continuing to monitor progress. In CPS for GEI, follow-up meetings are informal after the creation of the intervention. In BBT for GEI, the team meets again in formalized data review meetings to evaluate the plan by discussing progress monitoring data and revisiting the primary area of concern if the intervention needs to be modified (Blumberg Center, 2006).

BBT for GEI is more consistent with the best practices for problem-solving intervention teams when compared to the CPS for GEI, because it includes specifications for conducting the problem analysis and plan evaluation stages of the behavioral consultation model described in Allen and Graden (2002). Because of the emphasis placed on the use of data collection, progress monitoring and plan evaluation, the intervention plans created by the teams using BBT for GEI should be more robust than those created by the teams using CPS for GEI. The BBT for GEI process also includes a measure of implementation integrity to verify that teams are implementing the process as intended.

Effectiveness of Problem-Solving Teams

Previously, researchers have used five methods to evaluate the success of problem-solving teams. Sindaler et al. (1992) identified five measures that are usually cited in the literature as indicators of effective problem-solving team practice. The five measures cited by Sindaler et al. include (a) reduction in referrals to special education, (b) improved student performance on academic or behavioral concerns, (c) altered teacher expectations or perceptions, (d) teacher/parent/student satisfaction, and (e) an improvement in educational practice or teacher skill. The models previously described have all been researched in relationship to one or more of the five measures cited by Sindaler et al. and will be discussed in detail in the following sections.

Referral and Placement

One of the most commonly used methods for researching problem-solving teams is to evaluate the rates of referral and placement into special education before and after interventions are implemented. Chalfant and Pysh (1989) found that, when implemented, the TAT model resulted in appropriate referrals to special education 78% to 100% of the time, while inappropriate referrals decreased by 63% from the preimplementation year. A similar study conducted with MATs revealed special education referral rates of 8% and 13% for short- and long-implementation conditions, respectively, when compared to a 50% rate of referral for students from a control group (Fuchs, Fuchs, & Bahr, 1990). In other words, the use of MATs resulted in approximately 85% fewer special education referrals than the control group. Bahr et al. (2006) conducted similar research regarding CPS for GEI. They found that teams trained in CPS for GEI affected a decrease in rates of unverifiable referrals (students who were assessed for special education but did not meet criteria) from 35% to 14%, while the control teams (not

trained in CPS for GEI) had a reduction in unverifiable referrals from only 37% to 30% over the same time period.

Subjective Ratings

Research on TATs measured student behavioral and academic functioning, team effectiveness, and progress towards goal attainment, but the research has been done through subjective reports and ratings. In a study measuring teacher satisfaction with TAT, researchers found that the group members attributed their success to variables of principal support, team attitudes and training, team performance, and faculty support or participation (Chalfant & Pysh, 1989). Similarly, research on MAT has been evaluated through, for the most part, subjective reports. In one such study, it was found that MAT techniques resulted in fewer problem behaviors of difficult to teach students, increasingly positive teacher attitudes, and lower special education referral rates when compared to the control group (Fuchs, Fuchs, Gilman, et al., 1990). Fuchs, Fuchs, Gilman, et al. included a study of long and short MAT implementation conditions. Teachers in both short and long conditions responded positively to questions about intervention feasibility and effectiveness ratings. In addition, teacher perceptions of students' behavior were more positive in short-implementation conditions with regard to attention problems and anxiety/withdrawal problems for preintervention to postintervention; however, the study also found that in some schools, the teams were never fully implemented because of the time needed to run the multidisciplinary teams. The format of the MATs was later changed to consist of just a consultant and consultee with more emphasis on validated interventions from which the consultee was required to choose (Fuchs, Fuchs, Gilman, et al., 1990).

Student Outcomes

Another means of evaluating problem-solving teams' interventions is by tracking observable behavior changes or academic performance changes in student outcomes. In the Fuchs, Fuchs, and Bahr (1990) study of short and long MAT implementation conditions, the observed frequencies of targeted problem behaviors of the experimental group improved so much that the experimental group became indistinguishable from students not participating in MAT in regard to problem behaviors observed. The majority of the research on problem-solving intervention teams is conducted through teacher and team member ratings and referral rates for special education. This is true despite the stated goal of these teams to improve student performance and behavior (Fuchs, Fuchs, & Bahr, 1990). More research regarding student outcomes is needed in order to identify the effects of problem-solving intervention teams on student performance and behavior. The current study examined student outcomes and referral and placement rates because these two evaluation methods closely evaluate the effects of problem-solving intervention teams on student performance and behavior, the overarching goal of problem-solving intervention teams.

Intervention Process and Content Variables

A study of problem solving intervention teams should include procedures for evaluating process and content variables of the team-designed interventions (Strickler, 2004). Process variables include all aspects of the intervention design, while content variables consist of the components of the intervention developed from the design. Researchers would have difficulty determining which components of the problem-solving intervention process contributed to the student outcomes without evaluating process and content variables. By evaluating process and content variables of intervention design and relating them to student outcome data, researchers

can make more precise judgments as to which intervention components contributed most to student outcomes. To evaluate the process variables of intervention, researchers have identified quality indicators of intervention design (Flugum & Reschly, 1994; Strickler; Telzrow et al., 2000; Upah, 1998). Researchers have also evaluated the extent to which the chosen components of intervention include strong research-based elements in order to assess content variables (Flugum & Reschly, 1994; Strickler, 2004; Telzrow et al., 2000; Upah, 1998).

Quality Indicators of Intervention Design

The manner in which intervention teams operate is an important source of information. Research shows its relation to special education referrals and placement, teacher satisfaction, and to limited extent student outcomes. A goal of problem-solving intervention teams is to provide assistance to teachers in intervening with behavioral and academic concerns in the classroom (Fuchs, Fuchs, & Bahr, 1990; Lane et al., 2004). Typically, teams accomplish this by utilizing the four stages of behavioral consultation outlined by Allen and Graden (2002) to create an intervention plan that will be conducted in the general education classroom (Bahr et al., 2006; Chalfant & Pysh, 1989; Fuchs, Fuchs, & Bahr, 1990; Graden et al., 1985). Intervention plans identify who will be conducting the intervention, what they will be doing, where the intervention will be implemented, and how the steps of the intervention will be completed (Tilly & Flugam, 1995). In essence, the intervention plan outlines how the team will conduct the four stages of behavioral consultation. Researchers have attempted to expand on these stages, creating a list of quality indicators that, if used in the intervention design process, will result in better student outcomes from the problem-solving intervention process. The following list of represents the summation of the quality indicators that have been empirically evaluated in the literature and formed the bases for the current study: (a) behavioral definition of the target behavior, (b)

baseline data, (c) systematic intervention plan, (d) treatment integrity, (e) graphing of intervention results, (f) direct comparison of postintervention performance to baseline data, (g) problem validation, (h) functional analysis, (i) goal setting, (j) measurement strategy, (k) decision-making plan, (l) formative evaluation, and (m) intervention content variables.

Behavioral Definition of the Target Behavior

The first indicator of a quality intervention included in Flugum and Reshley (1994) was a Behavioral Definition. A behavioral definition has been identified by many researchers as an important component to intervention success (Baer, Wolf, & Risley, 1968; Deno, 1995). According to Kazdin (1982), a behavioral definition must meet three criteria: objectivity, clarity, and completeness. To be objective, the definition must refer to observable characteristics of the behavior. In order for the definition to be clear, it must be stated unambiguously so that it could be read and understood without the possibility of misunderstanding. In order for the definition to be complete, the conditions for which responses or observed behavior are delineated into both examples and non-examples of the behavior must be explained (Kazdin, 1982).

Baseline Data

According to Kazdin (1982), baseline data provide information regarding the level of behavior before an intervention begins. Kazdin further explains that baseline data are used to predict the level of performance overtime if no intervention is implemented. To be an effective predictor of performance, the data must be stable or without trend (slope) and with little variability in performance.

Systematic Intervention Plan

Tilly and Flugum (1995) have suggested that an intervention plan should explain who is responsible for intervention implementation and insuring progress monitoring data is collected.

In addition to this, Tilly and Flugum suggest that the plan should specify what he or she will be doing as well as when, where, and how often the intervention components will be carried out. Batsche and Knoff (1995) suggest that an intervention plan should include the following four components: (a) assessment outcomes, (b) specific intervention strategies, (c) expected outcomes with criteria for success, and (d) personnel responsible for implementing the plan.

Treatment Integrity

Treatment integrity is defined as treatment or intervention adherence (Strickler, 2004). Telzrow (1995) suggested that treatment integrity is the degree to which the intervention is implemented as planned. According to Upah (1998), treatment integrity has been identified as an important component to include in the intervention process, but that it has been under researched in the behavioral literature.

Graphing of Intervention Results

Graphing of data from the intervention allows for the progress monitoring to be readily examined. Kazdin (1982) outlined several considerations for visual representation of graphed data including providing a descriptive title, understandable scale captions with appropriate units, and marked phase lines for treatment or phase changes (either planned or unplanned).

Direct Comparison of Postintervention Performance to Baseline Data

A direct comparison of postintervention performance to baseline data is a summative evaluation. According to Upah (1998), the purpose of a summative evaluation is to determine whether an intervention was successful or not. A summative evaluation should consist of several data points in order to demonstrate that the intervention was responsible for the measured change in behavior (Steege et al., 2002).

Problem Validation

Problem validation refers to providing support for the identified problem by determining the magnitude or intensity of the problem (Strickler, 2004). Problem validation requires a comparison between an individual's current performance and/or baseline data to some standard measure of performance (Tilly & Flugum, 1995). This standard is often obtained through district norms, benchmark assessments, classroom comparisons of functioning, or developmental norms (Tilly & Flugum, 1995; Upah, 1998).

Functional Analysis

Often, further analysis of the variables surrounding a problem is necessary in order to design interventions. Functional analysis is used to test hypotheses about the reasons for a behavior. These functions may not be readily revealed through observations of natural events (i.e., Antecedent-Behavior-Consequence analysis) and other procedures may be necessary to determine the function of behavior (Barnett, 2002). According to Upah (1998), it is necessary to gather information for the functional analysis from a variety of sources, such as a review of records, interview, observation, and testing, in order to be aware of possible alterable factors from a variety of domains. A Curriculum Based Evaluation or CBE is an effective method for conducting functional analyses of academic concerns. Howell, Kurns, and Antil (2002) have defined CBE as a thoughtful process of comparison and judgment that provides a decision-making framework for hypothesizing the assumed causes of a student's academic problems, validating these hypotheses, and linking the results to a teaching recommendation.

Goal Setting

Goal setting consists of clearly articulating or delineating the desired outcome level of the intervention. According to Upah and Tilly (2002), a goal statement should include the following

four components: (a) a time frame for when the expected progress is to be accomplished, (b) the condition or specific circumstances under which the behavior is to occur, (c) the behavioral description of the task to be performed, and (d) the criteria or standard of how well the behavior is to be performed. Goal setting also allows for procedures to evaluate goal attainment as an outcome measure (Cobb, 1995).

Measurement Strategy

A measurement strategy is identified as part of the collection of baseline data. Additionally, the overall method of collecting data should remain the same throughout the intervention process. According to Upah and Tilly (2002), the following five questions should be answered in order to ensure the integrity of the measurement strategy: (a) How will the data be collected? (b) What materials will be used to collect data? (c) In which settings will data be collected? (d) Who will be responsible for collecting data? and (e) When and how often will the data be collected?

Decision-Making Plan

A decision-making strategy for determining how decisions will be made consists of identifying a strategy for data collection, summarization, and evaluation (Ross, 1995). Tilly and Flugum (1995) identified four specific issues related to a decision-making plan that should be addressed: (a) frequency of data collection, (b) strategies to summarize data for the evaluation, (c) number of data points or amount of time that should pass before the data will be analyzed, and (d) guidelines for action when certain patterns of data occur.

Formative Evaluation

A formative evaluation measures the plans successfulness during implementation phase. A formative evaluation should be conducted throughout the implementation of an intervention so

modifications can be made as necessary to improve results. Kazdin's (1982) visual analysis criteria, change in mean, level, trend, and latency is an effective tool for conducting a formative evaluation. Following the decision rule, which is based on progress monitoring data, is another method of conducting formative evaluation (Upah & Tilly, 2002).

Intervention Content Variables

The previous quality indicators were each process variables related to intervention. Strickler (2004) added four content variables as quality indicators of intervention. The four content variables were combined into one indicator of strong intervention. The four content variables are (a) naturalistic intervention, (b) instruction, (c) practice, and (d) modeling. Naturalistic interventions are interventions that fit into existing routines or experiences of the student. Stokes and Baer (1977) reported that using naturalistic interventions facilitated the generalization of behavioral changes outside the intervention. Effective interventions include the key components of instructional practices including scaffolding, shaping, connecting to prior knowledge, constructing meaning, and motivation (Gettinger & Stoiber, 1999). Providing opportunities to perform and respond within the intervention is a critical component of practice (Barnett, 2002). Effective intervention provides sufficient opportunities to practice (Strickler, 2004). An additional component of effective intervention is modeling. Learning through observing models has been shown to be an effective method for acquiring knowledge (Bandura, 1986).

Studies of the Quality Indicators of Intervention Design

Four studies illustrate the progression of the quality indicators over time. Flugum and Reschly (1994) examined the relationship of six quality indicators and student outcomes. Since

Flugum and Reshly's initial six quality indicators, the number of indicators identified in the literature has increased. Table 1 summarizes the expanding quantity of quality indicators.

Table 1

Quality Indicators of Intervention Design for Four Studies

Flugum and Reschly, 1994	Tilly and Flugum, 1995	Upah and Tilly, 2002	Strickler, 2004
1. Behavioral definition	1. Behavioral definition	1. Behavioral definition	1. Behavioral definition
2. Baseline data	2. Baseline data	2. Baseline data	2. Baseline data
3. Systematic intervention plan	3. Problem validation	3. Problem validation	3. Problem validation
4. Treatment integrity	4. Functional analysis	4. Problem analyses	4. Problem analyses
5. Graphed data	5. Goal setting	5. Goal Setting	5. Goal Setting
6. Comparison of baseline and postintervention data	6. Intervention plan development	6. Intervention plan development	6. Intervention matching hypothesis
	7. Treatment integrity	7. Measurement strategy	7. Intervention plan development
	8. Progress monitoring	8. Decision-making plan	8. Measurement strategy
	9. Program evaluation	9. Progress monitoring	9. Decision-making plan
		10. Formative evaluation	10. Progress monitoring
		11. Treatment integrity	11. Formative evaluation
		12. Summative evaluation	12. Treatment integrity
			13. Summative evaluation
			14. Effective intervention content
			15. Social validity

The six indicators discussed by Flugum and Reschly included (a) behavioral definition of the target behavior, (b) a direct measure of the student's behavior in the natural setting prior to intervention implementation (baseline data), (c) a systematic intervention plan, (d)

implementation of the intervention as planned (treatment integrity), (e) graphing of intervention results, and (f) direct comparison of the student's postintervention performance to baseline data. Flugum and Reschly found that many of the interventions created by problem-solving teams did not include quality interventions. Despite this, the number of quality indicators included in an intervention significantly correlated with reports of improved behavior and student functioning.

As noted in Table 1, Tilly and Flugum (1995) added three indicators: (a) problem validation, (b) functional analysis, and (c) goal setting to the six indicators identified by Flugum and Reschly (1994). Using eight of these indicators (with problem solving excluded), Telzrow et al. (2000) examined the fidelity of problem-solving implementation by problem-solving intervention teams in 227 schools in relation to student outcomes. To measure fidelity of problem-solving implementation, Telzrow et al. created a case evaluation rubric to measure the teams' submitted case documentation. The case evaluation rubric consisted of Likert scale ratings from 1-5 for each of the eight quality intervention indicators measured. A rating of 1 on a quality indicator meant the indicator was not present, while a rating of 5 meant that the indicator was implemented fully or to the highest fidelity. The quality indicators that were implemented with the highest fidelity included Behavioral Definition of the Problem, and Clearly Identified Goal. Indicators that were implemented to the lowest fidelity include Hypothesized Reason for the Problem (and Treatment Integrity. Student outcomes were measured by examining the teams' Evaluation Team Report (ETR), which documents student progress through the use of a goal attainment scale. The goal attainment scale consisted of ratings of student progress toward the identified goal (1 = evidence of student regression; 3 = progress remained the same; and 5 = goal was achieved or exceeded). The relationship between the fidelity of implementation of quality indicators and student outcomes was analyzed. Six

quality indicators correlated significantly with ratings of student outcome: Behavioral Definition of Problem, Baseline Data, Clearly Identified Goal, Systematic Intervention Plan, Functional Analysis, and Comparison of Student Performance with Baseline.

Upah (1998) examined the nine indicators of quality intervention identified by Tilly and Flugum (1995) on several variables: (a) the effect training in designing and implementing interventions has on the quality of intervention, (b) the effect protocol-based documentation based on the nine quality indicators has on the quality of intervention, and (c) the relationship between the quality indicators of the intervention and the outcome of the intervention. Upah examined 145 cases provided by the Iowa Heartland Area Education Agency 11. Upah examined the first case file submitted by each participant group (the group of individuals submitting cases) to ensure they included adequate documentation of the intervention process. The participants were divided into one of three treatment groups. Treatment group one received no intervention. Treatment group two received the protocol for documenting activities, while the third group received the protocol and training in designing and implementing intervention and a follow-up meeting after training. After the treatment phase of the research, each treatment group submitted new cases for comparison purposes.

To evaluate the quality indicators within the cases submitted, the nine quality indicators presented by Tilly and Flugum (1995) were coded for presence (present or not) and level of quality (coded on 1-5 scale) with 1 being not present and 5 being fully present. Upah (1998) found that the use of a protocol for documenting intervention design activities increased the fidelity of implementation to between 70% and 100%. Including the training to the protocol treatment also increased fidelity of implementation between 76% and 100%. Thus, the addition of protocol documentation and training increased the implementation of the quality indicators of

intervention. The quality indicators that were implemented to a high degree (4.00 or higher) were problem validation, progress monitoring, and program evaluation.

To examine student outcomes, Upah (1998) used teacher and practitioner ratings, expert ratings, and a visual analysis of documented change in performance, based on Kazdin's (1982) criteria for visual analysis. The criteria include change in mean, change in level, change in trend, and latency of change. Each measurement was coded using scores of 1-4 with a score of 1 being the least desirable score and 4 being the most desirable score. Baseline results for student outcomes presented by Upah produced outcomes ranging from 1.18 to 2.19. For the protocol only phase, ratings ranged from 1.44 to 2.89. The training plus protocol phase and the follow-up and protocol phase had ratings ranging from 1.75 to 3.14. Thus, the use of a protocol for documentation and training for design and implementation resulted in improved student outcomes when compared to baseline. Additionally, a significant positive relationship was found between the quality of intervention and student outcomes (ranging from .21 to .50). These results suggest that when interventions contain the nine indicators of quality intervention to a high level of fidelity, it is more likely the interventions will lead to positive student outcomes.

According to Upah and Tilly (2002), the nine indicators of quality intervention presented by Tilly and Flugum (1995) can be further broken down to improve clarity and isolate specific indicators. To the nine quality indicators (Tilly & Flugum, 1995), Upah and Tilly added measurement strategy decision-making plans and formative evaluation as noted in Table 1. Flugum and Reschly (1995), Telzrow et al. (2000), and Upah (1998) have found that schools are not implementing the quality indicators of intervention to a high degree. Upah and Tilly proposed a process of self-monitoring of quality intervention implementation to increase the overall quality of intervention services for students. The self-monitoring tool created by Upah

and Tilly is an innovation configuration. An innovation configuration (Hord et al.,1987) is a tool developed to assist teachers and help in defining the components of a program by describing in clear operational terms the variations in practice that may occur as practitioners apply new skills. In creating the innovation configuration for quality indicators, Upah and Tilly have identified the different levels of implementation for each quality indicator. The innovation configuration organizes the now 12 quality indicators of intervention into the four behavioral consultation stages (Bergan, 1977). For each quality indicator, the innovation configuration consists of five levels of implementation presented in a Likert type, ordinal scale with a rating of 5 indicating a best practice application and a rating of 1 indicating no application of the quality indicator.

Strickler (2004) used the innovation configuration developed by Upah and Tilly (2002) to examine how well interventions designed by problem-solving intervention teams participating in a specific regional training effort align with the quality indicators of intervention. To do this, Strickler created a quality intervention worksheet based, in part, on the innovation configuration developed by Upah and Tilly. Strickler's quality intervention worksheet included the 12 items from Upah and Tilly's innovation configuration as well as three additional indicators of quality intervention, including intervention/hypothesis, social validity, and intervention content. Again, Table 1 summarizes the progression of quality indicators from Flugum and Reschly (1994) to Strickler.

The quality intervention worksheet was developed through a three-stage procedure (Strickler, 2004). The first stage consisted of developing scaled items using a literature review of relevant research on intervention design. The second stage consisted of an expert panel review of the scaled items. The expert panel was comprised of two school psychology faculty members from the University of Cincinnati, two consultants from the Southwest Ohio Special

Education Regional Resource Center, and two practicing school psychologists. The panel was asked to rate the 15 quality indicators created in the first step of development on the "extent to which the indicators were discrete, included, and adequately represented on the worksheet" (Strickler, 2004, p. 40). The panel was then asked to review the appropriateness and fit of each quality indicator with the innovation configuration from Upah and Tilly (2002). The panel made recommendations regarding the clarity of the worksheet. The third stage of development for the quality intervention worksheet consisted of a pilot study to test the viability of the worksheet at assessing the degree of implementation of quality intervention and to identify any necessary scoring changes prior to the implementation of the worksheet as a measure (Strickler, 2004). The pilot study consisted of three independent raters using the quality intervention worksheet to rate five cases from the sample. Interrater reliability was calculated on two occasions, one after the raters reviewed three cases and one after the fifth review. Interrater reliability following the first review session ranged from 50-90%. Following the final two independent ratings, the interrater reliability was above 80%.

Strickler (2004) rated, in total, 51 cases submitted by intervention teams using the quality intervention worksheet. Twenty-five percent of those cases were coded for reliability resulting in a mean reliability of 84% (SD = 6.05). Out of the 15 quality indicators measured, the mean number of indicators present in the cases was 12.81 (SD = 1.76). The teams included the following quality indicators 100% of the time: (a) behavioral definition of target behavior, (b) goal setting, (c) intervention plan development, (d) effective intervention content, and (e) progress monitoring. Social validity (defined here as the acceptability and social meaningfulness of a target behavior, intervention, and outcomes to those involved with the plan) was found in only one half of the case files. Indicators that were implemented to a higher degree based on

quality intervention worksheet scores (1–5) included (a) operational definition of target behavior, (b) a measurement plan, and (c) effective intervention content. Treatment integrity and social validity were the lowest rated quality indicators. A significant correlation was found between the number of indicators present and degree of implementation. Thus, simply by implementing the quality indicators, it is suggested that teams usually do so to a high degree. A goal attainment scale was used to evaluate whether or not a relationship existed between the quality indicators and student outcomes. No significant correlation was found between the number of indicators and student outcomes, but a statistically significant correlation was found between goal attainment scale and ratings on the innovation configuration. This was however, only a modest correlation.

CHAPTER 3

METHOD

Schools across Indiana received training in the BBT for GEI process. Each of those teams was asked to submit to the Blumberg Center documentation of activities for actual students who were targeted by their problem-solving teams for intervention. The documentation provided by the problem-solving teams for each student became the source from which the participants for this study were found. For inclusion in the study, 15 schools that submitted at least five student cases were randomly selected from the total number of schools that submitted cases to the Blumberg Center. From each of the randomly selected schools, five students were randomly selected to have their cases included in the current study. Therefore, 75 unique students had their cases selected. A "case" consisted of a problem-solving team identified goal for a referred student and the accompanying documentation of the team's activities related to that goal. For those referred students who had multiple goals, each goal for that student was considered an independent case as long as each goal was uniquely addressed in the accompanying documentation. Therefore, every goal identified and addressed from the randomly selected students was included in this study as a unique case. This presented independence problems for the study, but was necessary to differentiate between problem-solving team activities for particular students. Of the 75 unique students whose cases were included in the study, 57 had only one goal, 15 students had two goals, two students had three

goals, and one student had four goals. A total of 97 cases were identified from the 75 students. The cases were comprised primarily of general education teacher referred students who had presenting concerns in either academic or behavior domains. The accompanying documentation of each case included the following: background/referral information, intervention/action plans, and team accomplishment sheets. The background/referral form included descriptive information such as age, sex, and ethnicity. It also included previous assessment and intervention information as well as personal information like attendance, medical, and discipline records. The intervention/action plan was a form for documenting the intervention plan and progress monitoring. It included a table with the following headings under which the teams could input their plan: (a) Intervention, (b) Materials, (c) Frequency, (d) Duration, (e) Target Start Date, (f) Baseline Data, (g) Actual Implementation Date, and (h) Progress Monitoring Data. The team accomplishment sheet is a form on which teams can document the outcomes of the intervention development and implementation including: (a) decisions obtained from initial meeting, (b) progress monitoring data, and (c) decisions for special education referral.

Case Demographics

The 97 cases that were included in this study represented schools from across Indiana. Schools were not required to submit school identifying information, so the specific school that had student cases included in the study is unknown. Of the 97 cases included in the study, which represent the problem-solving team identified goals of 75 students, all had intervention/action plans and background information/referral information forms. Nine of the cases had no team accomplishment sheet. On each intervention/action plan, the number of data review or progress monitoring meetings was indicated. Twenty of the included cases met only one time at the initial intervention planning meeting. These 20 cases did not submit any progress monitoring data.

Fifty-one of the cases indicated that two meetings were held, the initial planning meeting and one progress monitoring meeting/plan evaluation meeting. Many of these teams included information that progress monitoring data was collected several times between the two meetings so no assumptions about the quality of the intervention implementation or progress monitoring data can be made based on the number of formal meetings alone. Twenty-six cases had data indicating that more than two team meetings were held.

Of the 75 unique students that had cases represented in this study, 47 were male, 25 were female, and three did not specify gender. Twelve percent of these students were identified as African American, 52% as Caucasian, 27% as Hispanic, 3% were identified as other, and fewer than 7% identified no ethnicity. Among the 75 students, 13 % were kindergartners, 35% were first graders, 23% were second graders, 12 % were third graders, 8% were fourth graders, 4% were fifth graders, and 1% sixth graders. The grade levels for 4% of the students were not identified. For the 97 cases that were obtained from these students, a primary area of concern was noted. Table 2 summarizes the primary area of concern for each case.

Table 2

Primary Area of Concern for Each Case

Area of concern	Percentage
Reading	62.90
Written Expression	1.03
Oral Expression	1.03
Listening Comprehension	1.03
Math	13.40
Aggression	2.06
Impulse Control / Inattention	3.09
Social Skill	1.03
General Academics	6.19
Coping Skills	3.09
Other / Undefined	5.15

On the team accomplishment sheet, teams were to report if the student was still in the GEI process and whether he or she was referred for special education testing or not. If the student was referred for special education testing, whether the student was found eligible or not was to be noted. Based on this information, the cases were divided into four student groups: (a) referred but not eligible for special education, (b) referred and found eligible for special education, (c) still in the problem-solving process, and (d) no longer in the process but not referred. Thirty-three cases did not provide enough information on the team accomplishment sheet to be placed into one of these four groups. A fifth student group, insufficient information for student group placement, was created to account for these cases. Table 3 summarizes the number of cases that fell into each of these groups.

Table 3

Descriptive Statistics of Student Groups for Included Cases (N = 97)

Group	*N*
Referred but not eligible	05
Referred and found eligible	13
Still in the process	33
No longer in process	13
Insufficient information	33

Materials

Each case submitted by the problem-solving teams was rated using the BBT for GEI Action Plan Evaluation Worksheet (Appendix A). The BBT for GEI Action Plan Evaluation Worksheet is a tool expressly aligned to the BBT for GEI process designed to collect data regarding implementation of quality indicators of intervention design and student outcomes of BBT for GEI teams. The worksheet is adapted from the quality intervention worksheet created by Strickler (2004), which represents the most current accumulation of quality indicators of

intervention design presented in a measurable innovation configuration. The BBT for GEI Action Plan Evaluation Worksheet collects information on 13 of the quality indicators of intervention design noted by Strickler. The indicators of Social Validity and Intervention Matching Hypothesis were not included in the worksheet. Strickler rated Social Validity from surveys administered by the Ohio Department of Education. Similar surveys were not administered to any of the BBT for GEI teams and therefore obtaining ratings on Social Validity was not possible. The Intervention Matching Hypothesis indicator was removed because it was noted by Strickler to overlap significantly with the Problem Analysis indicator due to both addressing the creation of a hypothesis for the referral problem function. The BBT for GEI Action Plan Evaluation Worksheet contains Likert scales from1-5 on adherence and presence of each indicator. The criteria for the Likert scale ratings are included for each quality indicator. After the completion of the innovative configurations, the rater added a narrative below the rating to explain why the configuration received the rating it did. This was done to provide more in-depth information regarding why teams received the ratings they did.

In addition to information regarding the implementation of the 13 quality indicators, the Evaluation Worksheet provides demographic information for the case (e.g., grade, sex, and ethnic identity). The Worksheet also collects student outcome data in the form of two innovation configurations, Goal Attainment Scaling, and Student Measured Performance. These outcome variables were rated the same way as the 13 quality indicators of intervention design, on Likert scales from 1-5. Information on the student's current classification in the problem-solving process (e.g., still in the process and referred for special education testing) was another form of student outcome data collected on the Evaluation Worksheet. Finally, the Evaluation Worksheet provided a summarization table that separates the quality indicators by the components of the

problem solving model they represent: (a) problem identification, (b) problem analysis, (c) plan development, (d) plan implementation, and (e) plan evaluation. Table 4 details how the quality indicators make up the behavioral consultative components.

Table 4

Behavioral Consultative Components and Matching Quality Indicators

Problem Identification	Problem Analysis	Plan Development	Plan Implementation	Plan Evaluation
1. Behavioral Definition	1. Problem Validation	1. Goal Setting	1. Measurement Strategy	1. Formative Evaluation
2. Baseline Data	2. Problem Analysis	2. Intervention Plan Development	2. Decision-Making Plan	2. Treatment Integrity
		3. Strength of Intervention	3. Progress Monitoring	3. Summative Evaluation

Each BBT for GEI team was asked to complete the Intervention Implementation Fidelity Checklist (Appendix B) for each intervention implemented during the problem-solving process. On the checklist, teams were asked to write the percentage of time during each week of implementation the following five domains were carried out as designed: (a) implemented as defined, (b) frequency, (c) duration, (d) interventionist, and (e) data collection. The teams were then asked to return the Fidelity Checklist to the Blumberg Center along with the other documentation of activities. The Intervention Implementation Fidelity Checklist is a form for teams to document the extent to which the created interventions are being implemented as they were intended to be. The checklist was used to make innovation configuration ratings on the Treatment Integrity indicator. Ratings on the Treatment Integrity indicator were based on teams' combined summary percentages of implementation as reported on the Intervention

Implementation Fidelity Checklist. The higher the percentage the teams reported implementation on the Fidelity Checklist, the higher the rating.

Procedures

Each submitted case received a case number and any student identifying information was removed using a previously agreed upon method between the researcher and his committee. The Blumberg Center created photocopies of the selected cases with all identifying information blacked out. These blacked out cases were assigned case numbers. These blacked out cases were then reviewed and a BBT for GEI Action Plan Evaluation Worksheet was completed by the researcher for each.

Interrater Reliability

Two graduate students in school psychology at Indiana State University trained in school based consultation were recruited to independently rate 5% of the submitted cases on the BBT for GEI Action Plan Evaluation Worksheet to calculate interrater reliability. Interrater reliability refers to, "the extent to which independent observers, or judges, agree on a rating of a specific behavior or rating" (Burry-Stock, Shaw, Laurie, & Chissom, 1996, p. 254). The graduate students were trained in using the BBT for GEI Action Plan Evaluation Worksheet with the use of scored examples. They were also provided a reference key which they could consult during their analysis (Appendix C).

For measuring interrater reliability, rater agreement indices (RAIs) were calculated for the 13 quality intervention indicators BBT for GEI Action Plan Evaluation Worksheet individually and combined. RAIs have been recommended to assess the degree of agreement when ratings are derived from, "any kind of assessment material such as observations, portfolios, and interviews" (Burry-Stock et al., 1996, p. 260). RAIs are presented as a set of detailed

formulae to calculate the extent to which independent raters agree on a set of observations or variables. RAIs are calculated by using a series of algebraic formulae described in detail in Burry-Stock et al. (1996). RAI scores range from 0-1. The closer a RAI is to 1, the more the raters are in agreement. A score of 1 indicates that the two ratings or observations are identical.

RAI were calculated for rater agreement between the two graduate students and the researcher. The overall RAI, the combination of agreement indexes on all the quality indicators and the two outcome variables, was .87. The RAIs for individual indicators ranged from a low of .78 for the Problem Validation indicator to a high of 1.00 for Summative Evaluation. Table 5 includes the complete list of the RAIs for each quality indicator and the two outcome variables.

Table 5

Quality Indicator Reliability Agreement Indices (RAIs)

Quality Indicator	RAI
Behavioral Definition	.87
Baseline Data	.85
Problem Validation	.78
Problem Analysis	.97
Goal Setting	.83
Intervention Plan Development	.78
Strength of Intervention	.82
Measurement Strategy/Progress Monitoring Plan	.82
Decision Making Plan	.87
Progress Monitoring	.90
Treatment Integrity	.97
Summative Evaluation	1.00
Goal Attainment Scaling	.87
Student Measured Performance	.86

For the Problem Validation indicator, the rater had to make a decision if the data validating the problem statement included a comparison between the student's performance and local or national benchmarks or norms. Often student performance data was quantified and objective, but it was difficult to determine if the quantified data was compared to any particular norm (e.g.,

teams would report a percentage on an academic skills test but would not indicate what test was given or whether the test was normed). Because of this, the rater sometimes had to make an assumption on the nature of the test or rate the indicator as not meeting the normed data criterion. This likely contributed to the discrepancies on this RAI. The obtained RAIs suggest a high level of reliability for the ratings of the quality indicators.

Data Analysis

The first research question asked how well BBT for GEI teams' practices aligned with the quality indicators of intervention. The question was analyzed through descriptive statistics based on the quality indicators and student outcomes. To provide a more in-depth analysis of BBT for GEI teams' practices, each narrative was assessed for the presence of themes related to the ratings. These themes were further analyzed for overall common themes across the cases and quality indicators. The overall themes that emerged provided information as to why each innovation configuration received the rating it was given. These themes provided qualitative information regarding the specific activities of the BBT for GEI teams.

The second research question, what effects do the BBT for GEI teams have on student outcomes, was analyzed using two multiple regression analyses. For each regression analysis, the predictor variables were the 13 indicators of quality intervention design. The criterion variables were the two outcome variables: Goal Attainment Scaling and Student Measured Performance.

The third research question asked whether significant differences exist between the quality indicators of the interventions created by BBT for GEI teams and the outcomes for the student groups? The student groups included students referred but not identified for special education, students referred and found eligible for special education, students no longer in the

problem-solving process due to meeting goals, and students no longer in the problem-solving process due to unknown reasons. The student groups were analyzed using an analysis of variance (ANOVA) comparing each student group to examine if differences exist on the outcome measure. Students remaining in the problem-solving process were not included in this analysis because it was not known in which post-process group these students belonged. The independent variable of the ANOVA was the student group (referred but not identified, referred and identified, or the two no longer in GEI groups). The dependent variable for the ANOVA was the overall student outcomes score obtained by summing the means of the two outcome scores, Goal Attainment Scaling, and Student Measured Performance.

CHAPTER 4

RESULTS

The results of this study are presented in the following order: descriptives of the referral cases, ratings of quality indicators (individually and together), predictors of student outcomes, and differences between student groups. A total of 75 students had cases analyzed. Ninety-Seven cases were analyzed in all.

Research Question One

The first research question, was what extent do teams' practices align with quality indicators. To measure this, each case's quality indicators were rated on the degree of implementation. A quality indicator that was implemented to a low degree, or possibly not addressed, received a low rating. As the quality indicator implementation increased in quality, the rating of the indicator was higher. In this manner, all of the indicators in each case received ratings between 1 (low) to 5 (high). The more key elements in an indicator that were present, the higher the rating. The fewer key elements present in an indicator, the lower the rating. For example, on the Behavioral Definition indicator a rating of 5 is only given if the case includes a behavioral definition of the problem behavior that is clear, objective, and complete. A rating of 4 is given for cases including two of these criteria, and a rating of 3 is given if only one criterion is met. A rating of 2 is given for cases only identifying the problem behavior in general terms while a rating of 1 is given if no definition of the problem behavior is given. The BBT for GEI

Action Plan Evaluation Worksheet (Appendix A) delineates the criteria for each of the ratings, 1-5, for each quality indicator. The average indicators ranged from a low rating of 1.44 to a high rating of 3.64 (See Table 6).

Table 6

Quality Indicator Mean Ratings

Quality Indicator	*M*	*SD*
Behavioral Definition	2.68	1.09
Baseline Data	3.12	0.97
Problem Validation	3.57	1.43
Problem Analysis	3.41	0.81
Goal Setting	3.42	1.43
Intervention Plan Development	3.64	0.99
Strength of Intervention	3.04	0.97
Measurement Strategy/Progress Monitoring Plan	2.53	1.29
Decision Making Plan	2.22	0.93
Progress Monitoring	2.28	1.28
Formative Evaluation	1.46	0.97
Treatment Integrity	1.97	1.64
Summative Evaluation	2.30	1.42

Quality Indicators Rated 3.5 or Higher

Of the 13 quality indicators, two received ratings of 3.5 or higher: Intervention Plan Development ($M = 3.64$) and Problem Validation ($M = 3.57$). For a case to receive a 5 point rating on the Intervention Plan Development indicator, it needed to include the following key elements: (a) a description of the intervention procedures, (b) intervention materials, (c) when the intervention would be implemented, (d) where the intervention would be implemented, and (e) who would be responsible for implementing the intervention. When analyzing the narratives that accompanied each of the ratings for the Intervention Plan Development, the element that was most often missing (67%) from the indicator was a statement of where the intervention would take place. This is likely due to the fact that the BBT for GEI Action Plans did not

include an explicit place to list where the intervention was to be conducted. Some teams (18%) did not include enough information on the procedures to receive credit. These teams simply wrote a generic or vague description of the planned intervention. Of the 11 cases in the study dealing with behavioral issues, eight provided only a generic or vague description of the planned intervention (e.g., the counselor will conduct a lesson on personal space). When teams utilized interventions with standard procedures (e.g., participate in Reading Recovery or the computer reading program Odyssey), they only had to include the name of the standard procedure to receive full credit for the information on the procedures criterion. Teams rarely (12%) received lower ratings for not including the elements related to who was responsible and what materials were to be used. Both of these elements had explicit locations for documentation on the action plans.

To receive a rating of 5 on the Problem Validation indicator, the team was required to quantify the magnitude of the problem based on a comparison between the student's performance and local or national norms or benchmarks. Cases that did not include a comparison to a norm or benchmark could receive a score no higher than a 3. Cases that used objective, quantitative data were given higher ratings than cases with subjective data. A review of the narratives for Problem Validation revealed that almost all of the cases (91%) included some data to support and validate the identified problem. The background information form for each case included much of the data for problem validation (e.g., ratings of severity, teacher observations, and state-wide testing results). Sixty-seven percent of the cases included at least some objective data such as achievement test scores. Nearly all of the cases (90%) included subjective data like statements of magnitude, observational data, and teacher ratings of severity. Some of this data clearly met the criteria of being normed or benchmarked, such as state-wide

academic assessments. Thirty-seven cases included this normed data. Some data that appeared normed could not be scored as such because it was unclear whether the validating data was in fact normed. One example of seemingly normed data that could not be scored as such includes reading fluency scores presented as words-per-minute without a reference to a norm or benchmark for the words-per-minute score. A total of 21 cases validated the problem in this way. These teams used quantified data between the student's performance and some standard, but they did not include any explanation or interpretation of the standard (i.e., the team did not explain what the standard means or where it came from). It is possible that the team knew what normative data was being used and did not include this information in the background data because all the interested parties already know the bases of the norm.

Quality Indicators Rated 3 to 3.5

Four indicators received ratings from 3 to 3.5. These included the following: Goal Setting (M = 3.42), Problem Analysis (M = 3.41), Baseline Data (M = 3.12), and Strength of Intervention (M = 3.04). The first of these, Goal Setting required the following four components to be included: time frame, behavior, criterion (level of behavior/achievement expected), and condition (circumstances under which the behavior is to occur). To receive a score of 5, all of these components had to be present and the criteria of the goal must be based on a link between baseline data and the expectation or benchmark goal. In almost all of the cases reviewed, teams included a goal statement (89%). An overall theme for how teams were rated on this indicator was not found. Teams varied considerably on which components were included and which were not. Thirty-two cases included all of the components for Goal setting. Nineteen cases included three of the components, and 16 included two components. Eighteen cases included one of the four components. Teams that included objective baseline information and behavioral definitions

tended to include higher quality goals. Some teams that did include high quality behavioral definitions and baseline data, however, failed to match the stated goal to those quality indicators.

The second indicator rated between 3 and 3.5 was Problem Analysis. The Problem Analysis indicator represented the procedures teams used to identify all the possible contributing variables of the identified problem. Teams that analyzed more possible contributing variables received higher ratings on the quality indicator. In addition to analyzing possible contributing variables, a team must include a statement on their hypothesis of the function of the problem in order to receive a rating of 5. None of the teams rated included a hypothesis about the function of the behavior. Most teams (94%) did analyze personal factors (e.g., motor skills, language, and health history). This information was asked for on the Background Information form. Hypothesis about the function of the behavior was not asked for on the Background Information form. Teams also considered previous instruction, classroom environment, and in-depth behavior investigations to analyze the target problem. The consistency with which teams included their analysis of all of these factors is unknown. Teams did not document the ways in which the background information contributed to intervention development. This kind of documentation was not asked for on the BBT for GEI forms. Teams may have used the background information during the GEI meetings but one can only speculate on the extent to which they did. Anecdotally, it appeared many teams may have approached the Background Information forms as a sort of paperwork prerequisite to the GEI meetings. This speculation is based on the fact that the identified problems often did not seem to be linked to any background information. On these occasions, the background information was likely set aside once the team began creating interventions for the student. This is, of course, speculation because teams were not required to document how they used the background information.

The next indicator rated between 3 and 3.5 was Baseline Data. A quality Baseline Data indicator consisted of objective, measureable data collected multiple times prior to the intervention until a baseline trend was established. Teams that included baseline data and met this requirement received a rating of 5 on the Baseline Data quality indicator. Teams that included fewer data points or nonobjective data received lower ratings. Teams that did not report baseline data prior to the intervention being implemented received ratings of one. To receive a rating of 3 or higher, the teams had to include quantitative data. A review of the narratives for this indicator indicates that 76% of cases did include objective, quantitative data. Teams tended to only use one data point for the baseline information. Of the teams that used quantitative data, 47% included only one data point. Only nine cases used repeated measures to verify a baseline trend. Twenty-one cases did, however, use multiple measures to establish a baseline of performance. For instance, teams would report reading scores on several different tests to establish a baseline for a reading problem.

Strength of Intervention was the final indicator with an average rating between 3 and 3.5. The Strength of Intervention indicator measured team use of empirically supported strategies for intervention. The elements that make strong interventions include direct instruction, practice, modeling, positive contingencies, error correction, and naturalistic interventions. The more elements of a strong intervention that the teams included, the higher the rating they received. This rating was not based on implementation of these strategies, but on whether the proposed intervention included procedures that are empirically supported. Over three-quarters of the cases reviewed included at least one element of a strong intervention (84%). A review of the narratives for this indicator revealed that the strong element of intervention most commonly used was practice. Typically this would consist of additional practice with a skill (e.g., repeated

reading interventions, practice sounding out words, or practicing math operations). Often skill practice would be paired with direct instruction in the skill. Teams rarely (6%) paired practice or direct instruction with a reinforcement contingency. In fact, reinforcement strategies were typically only used for students with behavior problems, of which only 11 were included in this study. Cases that received ratings of one (12%), which indicates that no element of a strong intervention was used, received this rating because no intervention procedures were reported. Most of the interventions included in the cases were directly related to the stated goal (i.e., teams linked the intervention to the identified weaknesses).

Quality Indicators Rated 2 to 3

Five indicators received ratings from 2 to 3. The first of these was Behavioral Definition ($M = 2.68$). For a case to receive a rating of 5 on this indicator, the definition of the problem behavior had to be objective, clear, and complete. Cases received lower ratings if they did not meet all of these criteria. A score of 2 was given if the problem behavior was defined in general terms only (e.g., reading, behavior, or mathematics). A case received a score of one if it included no definition of the problem behavior. Only 7% of the cases rated included no behavioral definition. Just over half of the cases (51%) included only a general statement of the problem behavior. Forty-one cases received ratings of 3 or higher. The review of these narratives suggested that these cases included more narrowly defined academic goals than those using only general statements (e.g., defined the problem as reading fluency as opposed to simply identifying it as reading). The cases that received the highest ratings referenced specific measures of target behaviors. For instance, a team would receive a higher rating if they defined a reading fluency problem by stating it in terms of performance on an oral reading fluency test.

The oral reading fluency test clearly specifies the procedures for measuring the behavior and eliminates ambiguity in the definition.

The second indicator with a rating between 2 and 3 was Measurement Strategy/Progress Monitoring ($M = 2.53$). This indicator rated the teams' plans to measure the strategy or intervention implemented. This rating does not reflect the teams' actual implementation of the measurement strategy. For a team to receive a rating of 5 on the Measurement Strategy/Progress Monitoring indicator, it had to address set criteria concerning measurement strategy: how to measure progress, what measurement tool was used, where will the measurement take place, who is responsible for obtaining the measurements, and when will the measurements of progress be taken. Two-thirds of the cases rated included more than one of these components of progress monitoring (66%). These cases indicated what measurement tool would be used, who would be responsible, and when the progress monitoring will be done with equal regularity. Only five cases indicated how the progress would be monitored or where the progress monitoring would take place. These components may very well have been addressed by the teams and not included in the Action Plan. In fact, credit was given if the case implied where the progress monitoring was going to take place (e.g., during intervention class or at home for home-based interventions) or if the plan suggested some standard procedures like curriculum based assessment. Nonetheless, teams' progress monitoring plans needed to explicitly state where the progress monitoring was to be conducted because progress monitoring could have been carried out anywhere, and not necessarily in the same setting in which the intervention was conducted.

Summative Evaluation ($M = 2.30$) was the third quality indicator rated between 2 and 3. The Summative Evaluation indicator rated teams' outcome decisions after the implementation of the intervention. This rating does not measure the level of progress on the interventions; rather,

it measures the extent to which the teams used progress monitoring data to make their decisions. A rating of 5 on this indicator means that the team based its outcome decision on the progress monitoring data. A rating of 4 was given for those teams who used minimal data (e.g., pre and post data). A rating of 3 was given if subjective data was used, and a rating of 2 was given if outcome decisions were made but no data was used. A rating of 1 was given for cases that included no summative evaluation. Of the cases that did include a summative evaluation, over 30% (n = 24) did not indicate that any data was used to make outcome decisions. The review of the narratives for these cases revealed that many of these teams reported student outcomes but no decision for future student programming were made. Eighteen percent referenced only subjective data for the basis of the decision making. Typically the subjective data was in the form of a statement of progress (e.g., student improved performance). Over 20% of cases did base their outcome decisions on objective data, usually in the form of preintervention and postintervention progress monitoring data. Teams that used objective data for decision making represented 20% of the overall sample and 26% of the cases that met more than one time. Thirty-nine cases included no summative evaluation. This includes the 20 cases that included only one GEI meeting. Teams that met only once did not have an opportunity to evaluate intervention successfulness. When those teams that met only one time are excluded from the case ratings, Summative Evaluation is rated higher (n = 77, M = 2.62).

The fourth indicator rated between 2 and 3 was Progress Monitoring (M = 2.28). Ratings on the Progress Monitoring indicator reflect the extent to which the teams collected consistent data to measure the implemented intervention's effectiveness. For a rating of 5 on this indicator, teams were required to document that data was collected consistently using the same measurement tool and chart or graph that progress monitoring data for visual analysis. If

progress monitoring data was not collected at all, the case was given a rating of 1. Like the Summative Evaluation indicator, those cases that met only one time did not have the opportunity to collect progress monitoring data. For the cases that met more than once (n = 77), 87% of the cases included some form of progress monitoring data. Over 50% of these cases, however, included only preintervention/postintervention data. The narrative analysis indicated that the teams who did conduct more than just pre- and post data tended to be inconsistent in their data collection.

The fifth and final quality indicator that was rated between 2 and 3 was Decision Making Plan (M = 2.22). The Decision Making Plan indicator measures the teams' planning on how decisions were to be made regarding the cases (i.e., what to do if progress is not made or when the team will meet to review student progress). For teams to receive a rating of 5, the case had to contain documentation of the following: how frequently the data was to be collected; what strategies were going to be used to summarize the data for evaluation, how much progress monitoring would take place before making a decision or how much time will occur before analysis; and what actions will be taken based on the intervention data. Only 6% of the cases rated included at least three of these four criteria. Twenty-five percent of the cases included no decision making plan. The review of the narratives of this indicator revealed that many of the teams did indicate when they would meet again to review the progress monitoring data. This tended to be scheduled a few months in advance. Teams indicated what actions may be taken based on their data review on only three cases. Teams also rarely (3%) planned for how to determine when action plans would be successful at improving student behavior or achievement.

Quality Indicators Rated Below 2

Two quality indicators received ratings lower than 2. The first of these is Treatment Integrity (M = 1.97). The Treatment Integrity rating measured the extent to which teams monitored their implementation of the interventions as they were designed to be implemented. Teams were given a worksheet, the Intervention Implementation Fidelity Checklist (Appendix B) by the Blumberg Center to track their treatment integrity. Ratings on the Treatment Integrity indicator were based on teams' combined summary percentages of implementation as reported on the Intervention Implementation Fidelity Checklist. Cases with at least 76% implementation received a rating of 5. Cases where no consideration for treatment integrity was given received a rating of one. Seventy-one percent of the cases rated included no report of treatment integrity. Of the 26 cases that included the Intervention Implementation Fidelity Checklist, 19 reported treatment integrity at the 76% or higher implementation rate. When the cases that met only once are excluded (n = 20), the average rating for the Treatment Integrity indicator was higher (M = 2.22). This indicates that the typical teams that met to review their cases implemented at least some treatment integrity procedures during the intervention.

The second indicator rated lower than 2 was Formative Evaluation (M = 1.46). The Formative Evaluation indicator rated the teams' use of some procedure to make changes to the intervention during the implementation phase of the intervention based on progress monitoring data. For teams to receive a rating of 5 on this indicator they were required to include some decision rule to follow with procedures for modifying or changing the intervention as needed during implementation. The higher the quality of the data used to make the formative evaluation, the higher the rating teams received on this quality indicator. The ratings on this indicator indicate that most teams, 77%, did not include any form of formative evaluation. Like

the Treatment Integrity indicator, when the cases that met only one time are excluded from the analysis, the indicator ratings were higher. When the cases that met only once are excluded (n = 20), the mean rating was 2.62. For these cases, the teams were more likely to include some form of formative evaluation. A review of the narratives for this indicator suggest that teams often made adjustments to the intervention during the implementation phase, but there was rarely a decision plan in place to determine when to make changes and what changes to make if progress was not being made.

Differences between Cases that Met Once and those that Met Twice

As noted earlier, ratings were higher on some indicators when the cases that met in the problem-solving process only once were excluded from the analysis. To determine if this difference was significant, independent samples t tests were conducted with those cases that met only once (n = 20) and those cases that met more than once (n = 77) for each of the 13 quality indicators. For the Summative Evaluation indictor, there was a significant effect for number of meetings, $t(-9.27) = 87.95$, $p < .001$, with those cases that met more than once receiving significantly higher ratings than those that met only once. The effect size was large (d = 1.08). The same was true for Treatment Integrity, $t(76) = -6.09$, $p < .001$, and Formative Evaluation, $t(76) = -4.86$, $p < .001$. The effect size was large for both Treatment Integrity (d = 0.98) and Formative Evaluation (d = 0.77). On both indicators, cases that met more than once received statistically significant higher ratings than those the met only once. These results were expected because teams that met only once could not possibly evaluate the interventions they created. Table 7 includes the comparison of the indicator rating averages for those cases that met only once and those that met more than once. Significant differences in quality indicator ratings were

found between the two case groups on eight of the 13 quality indicators. On each of these, the cases that met more than once received higher ratings than those that met only once.

Table 7

Indicator Rating Averages Including and Excluding Cases that Met Only Once

Quality Indicator	Mean Quality Indicator Ratings[a]				
	cases that met only once (n = 20)	cases that met more than once (n = 77)	*t*	*df*	*Cohen's D*
Behavioral Definition	1.95 (0.39)	2.87 (1.14)	-5.86**	87.24	1.08
Baseline Data	2.75 (0.79)	3.22 (1.00)	-2.25*	36.51	0.52
Problem Validation	2.50 (1.15)	3.84 (1.37)	-4.48**	34.44	1.06
Problem Analysis	2.95 (0.94)	3.53 (0.74)	-2.56*	25.31	0.69
Goal Setting	3.25 (1.60)	3.47 (1.49)	-7.00	36.97	0.14
Intervention Plan Development	3.55 (0.83)	3.66 (1.03)	-0.51	36.13	0.12
Strength of Intervention	3.10 (1.07)	3.03 (0.95)	0.28	27.19	0.07
Measurement Strategy/ Progress Monitoring Plan	2.05 (1.19)	2.65 (1.30)	-1.97	31.72	0.15
Decision Making Plan	2.25 (0.97)	2.21 (0.92)	0.18	28.65	0.04
Progress Monitoring	1.15 (0.49)	2.57 (1.26)	-7.87**	80.90	1.49
Formative Evaluation	1.00 (0.00)	1.58 (1.06)	-2.47**	76.00	0.77
Treatment Integrity	1.00 (0.00)	2.22 (1.76)	-6.09**	76.00	0.98
Summative Evaluation	1.05 (0.22)	2.62 (1.42)	-9.27**	87.95	1.55

[a]Standard Deviations appear in parentheses next to the means.
*p < .05. **p < .001.

Four of the 13 indicators, Progress Monitoring, Formative Evaluation, Treatment Integrity, and Summative Evaluation represent indicators that can only be done on follow up meetings. The other four indicators showing significance, Behavioral Definition, Baseline Data, Problem Validation, and Problem Analysis, represent tasks that are typically done on the first meeting. It may be that teams that do these indicators to a high degree are more likely to reconvene the meeting because they have put in place a plan that specifies procedures for follow up meetings.

These teams may also have been more likely to schedule a follow up meeting as part of the procedures followed in the first meeting.

Behavioral Consultation Components

The quality indicators represent pieces of each step of the behavioral consultation components. These include problem identification, problem analysis, plan development, plan implementation, and plan evaluation. Table 2 summarizes the quality indicators associated with each behavioral consultation component. The behavioral consultation component that the teams implemented with the greatest fidelity was problem analysis (M = 3.49, SD = .95). The teams implemented plan development (M = 3.37, SD = .80) with the second greatest fidelity, followed by problem identification (M = 2.90, SD = .80) and plan implementation (M = 2.34, SD = .89). Teamsperformed the plan evaluation component with the least amount of fidelity (M = 1.91, SD = .99). Contributing to this was the fact that many of the teams (21%) did not meet more than one time to discuss the case. These teams, therefore, had no opportunity to evaluate the plan they had put in place.

Research Question Two

The second research question was what effect does the process have on student outcomes? Two student outcome variables were measured on the Quality Indicators of Intervention Worksheet: Goal Attainment Scaling and Student's Measured Performance. The Goal Attainment Scaling variable was measured the same way as the quality intervention indicators. Likert ratings were assigned from 1-5 to rate the extent to which the case met the team identified goal for the student. For a case to be given a rating of 5, the student goal had to be met or exceeded in both time frame and criterion. A rating of 4 was assigned to cases that met or exceeded the goal criterion but whose time frame for meeting the goal was not met or not

established. A rating of 3 was given to teams that made progress towards the goal but that did not meet the goal. A rating of 2 was given to cases that had indicated no change from the baseline data toward the identified goal. A rating of 1 was assigned to cases that showed decreases from the baseline. A total of 72 cases indicated some measureable level of goal attainment. The average rating for Goal Attainment Scaling was 2.92 (*SD* = 1.03). Twenty percent of the cases met or exceeded their goal. Nearly 73% of the cases made at least some progress toward the established goal.

Ratings were assigned for Student's Measured Performance the same way as for Goal Attainment Scaling with Likert ratings from 1-5. Ratings were based on the visual analysis of documented change in performance based on Kazdin's (1982) criteria for visual analysis. The criteria include change in mean, change in level, change in trend, and latency of change. When there are more of these criteria apparent in the outcome data, then rating on the Student's Measured Performance indicator will be higher. Because many of the teams did not present their outcome data in charts, graphs or other visual forms, the ratings on this indicator were not solely based on visually presented data. Instead, the rater examined both visual data and numerical data for changes across the visual analysis criteria when possible. Forty-five percent of the cases included in this study did not report outcome data that could be analyzed on any of the visual analysis criteria. Of the 43 teams that did reported outcome information, 18 cases did so in terms of general statements of improvement only. This information could have been scored on the Goal Attainment Scaling indicator, but a visual analysis of these types of subjective data is not possible. For these reasons, the average rating for Student's Measured Performance (*M* = 1.93, SD = 1.05) was substantially lower than Goal Attainment Scaling ratings. Teams that did provide outcome data that could be visually analyzed typically received credit for changes in

level. Because most of the cases analyzed included only pre- and postintervention data, changes in latency (i.e., no improvement until implementation phase), mean, and trend could not be observed.

Process Effect on Attainment of Goals

A multiple regression analysis was conducted to develop a model for predicting student goal attainment from the quality indicators of intervention. The quality indicators of intervention are numerated on the Quality Indicators of Intervention Worksheet (see Appendix A). They represent the best overlap between the BBT for GEI process and the research identified best practices for conducting general education intervention. Because of this, the quality indicator ratings obtained on the Quality Indicators of Intervention Worksheet became the predictor variables. The criterion variable was the Goal Attainment Scaling (GAS) indicator rating. Cases for the students that met only one time during the GEI process were excluded from the analysis listwise because these cases did not report intervention implementation data and therefore did not receive GAS ratings. Twenty cases met only one time and an additional five cases did not receive GAS ratings. This left 72 cases for inclusion in the analysis.

The assumptions underlying the multiple regression residuals (errors in prediction) were analyzed. The regression assumptions include the following: independence (not correlated to one another), normality (normally distributed residuals with a mean of zero), and homogeneity of variance (same variance across scale). A visual analysis of the plot of residuals was conducted to investigate these assumptions. The plot of residuals appeared random and independent, meeting the assumption of independence (keeping in mind the variables were Likert type and the residuals reflected the whole number ratings by falling along whole number intervals on the residual plot). The residual plot also maintained a constant scatter from left to right across the

plot suggesting the assumption of homogeneity of variance was upheld. The assumption of linearity among the variables themselves (predictors are linearly related to the criterion) was also checked. Only two of the residuals fell outside the +2.0 and -2.0 vertical and horizontal axes on the residual plot. This suggests that the variables, GAS and the 13 predictor indicators, are linearly related and thus upholds the assumption of linearity. To assess the assumption of normality, a histogram and normal probability-plot were created. The histogram follows a bell-curve pattern closely, with a slight negative bias. The normal probability plot follows a diagonal prediction line very closely. The residuals have a mean of zero, and based on the histogram and normal probability plot, the assumption of normality is upheld.

Another assumption underlying multiple regression analyses is the assumption of no multicollinearity, or intercorrelations among the predictor variables (the 13 quality indicators). To measure multicolliniarity, the tolerance statistic was calculated. Tolerance indicates whether a predictor has a strong linear relationship with the other predictors. Tolerance ranges from 0-1. The closer to 1 the tolerance statistic is, the lower the risk of multicollinarity to the regression model. Menard (1995) suggests that values below .2 are worthy of concern for multicolliniarity. The tolerance statistics for the predictor variables for this study ranged from a low of .29 for Progress Monitoring to a high of .71 for Problem Analysis. The assumption of no multicolliniarity was upheld. All of the assumptions for multiple regression analysis were upheld.

A total of 72 cases were included in the analysis. Descriptive statistics for these cases were previously reported in Table 7. A simultaneous regression method was used in order include all of the indicators in the analysis. The linear combination of the quality indicators of intervention design ratings was significantly related to the Goal Attainment Scaling ratings,

$F(13, 58) = 2.85$, $p < .01$. The multiple correlation coefficient was .62. The proportion of variability in the Goal Attainment Scaling indicator that can be attributed to differences in Quality Indicator ratings is 39% ($R^2 = .39$). To determine how much variance in Goal Attainment Scaling would be accounted for if the model had been derived from the population from which the sample was taken, the Adjusted R square was determined. The Adjusted R square was .25, indicating that the derived model works better for the sample than would be expected in the actual population. The standard error of the estimate was .89. This index indicated how large the typical error was in predicting GAS from the Quality Indicator rating. For a successful regression model, the standard error of the estimate should be smaller than the standard deviation of the dependent variable. The standard deviation of the GAS, the dependent variable, was 1.03. This means that the observed errors in predicting GAS were smaller than the observed differences in the mean for GAS ratings suggesting a successful regression model.

A review of the partial regression coefficients for the 13 quality indicator ratings revealed only two indicators, Problem Analysis, $t(56) = 3.15$, $p = .003$, and Intervention Plan Development, $t(56) = -2.05$, $p = .045$, as being significant predictors of GAS when the impact of the other predictors were removed. Interestingly, the Intervention Plan Development indicator had a negative relationship with GAS ratings. This means that as ratings on the Intervention Plan Development indicator increase by one, ratings on the GAS decrease by .31 ($b = .31$) when the other indicators are held constant. It is conceivable that when teams do not report what interventions are being used they may have no real basis for determining whether or not the student achieved the goal set by the team. They may report higher levels of goal attainment because they do not have objective, quantified data to make informed judgments on the level of goal attainment. This may help in explaining how the Intervention Plan Development indicator

could be negatively related to GAS. One possible reason the Problem Analysis indicator was a significant predictor could be the overall emphasis the BBT for GEI process places on problem analysis activities. Teams that do problem analysis activities to a high degree are likely investing a significant amount of time in the process compared to teams that do poorly in problem analysis. It would make sense that these highly invested teams would obtain positive results and thus higher GAS ratings. Based on the model when all the other indicators are held constant, an increase in rating on Problem Analysis by 1 will result in a .53 (b = .53) increase of GAS rating. Table 8 includes the partial regression coefficients for each of the 13 quality indicators.

Table 8

Summary of Multiple Regression Analysis for Variables Predicting GAS (N = 72)

Variable	*B*	*SE B*	β	*t*	*p*
Behavioral Definition	-.004	.12	-.01	-0.03	.974
Baseline Data	.203	.16	.20	1.31	.195
Problem Validation	.138	.09	.18	1.47	.148
Problem Analysis	.527	.17	.38	3.15	.003*
Goal Setting	.063	.08	.09	0.76	.448
Intervention Plan Development	-.307	.15	-.31	-2.05	.045*
Strength of Intervention	.183	.14	.17	1.33	.188
Measurement Strategy/ Progress Monitoring Plan	-.133	.11	-.17	-1.27	.211
Decision Making Plan	-.009	.14	-.01	-0.07	.947
Progress Monitoring	.003	.16	.01	0.02	.983
Formative Evaluation	.180	.14	.19	1.32	.192
Treatment Integrity	.017	.08	.03	0.22	.829
Summative Evaluation	.113	.12	.15	0.95	.347

Note. R^2 = .39 for Model. Adjusted R^2 = .25. df = 58.
* $p < .05$.

Process Effect on Student Performance

A multiple regression analysis was conducted to determine the effect the BBT for GEI process has on the Student's Measured Performance (SMP). The quality indicators of

intervention were again used as the predictor variables. The rating on SMP was used as the criterion variable. Cases were excluded listwise, which again excluded those cases that met only one time from the analysis. The assumptions underlying the multiple regression were analyzed the same way they were for the GAS regression model. The plot of residuals was visually analyzed to test for independence and homogeneity of variance. The residual plot appeared random and independent. Each of the assumptions related to the residuals was upheld. The variance was consistent across the scale and all but two of the residuals fell between the +2.0 and -2.0 vertical and horizontal axes, therefore the assumption of linearity between the variables was also upheld. A review of the histogram of the residuals and the normal probability plot revealed normally distributed data suggesting the assumption of normality was also upheld. Tolerance statistics were calculated to assess the assumption of no multicollinearity. The tolerance statistic for each of the 13 predictor indicators was greater than .20. The tolerance statistics suggest the assumption of no multicolliearity was also upheld.

A simultaneous regression method was used in order include all of the indicators in the analysis. The linear combination of the quality indicators of intervention design ratings was significantly related to the Student Measured Performance, $F(13, 58) = 5.46$, $p < .01$. The multiple correlation coefficient was .74. The percentage of variability in the SMP indicator that can be attributed to differences in Quality Indicator ratings is 55% ($R^2 = .55$). To determine how much variance in SMP would be accounted for if the model had been derived from the population from which the sample was taken, the Adjusted R square was determined. The Adjusted R square was .45, indicating that the model maintains stable predictability outside the current sample. The standard error of the estimate was .78. This index indicated how large the typical error was in predicting SMP from the Quality Indicator rating. As noted earlier, for a

successful regression model, the standard error of the estimate should be smaller than the standard deviation of the dependent variable. The standard deviation of the SMP, the dependent variable, was 1.05. This means that the observed errors in predicting SMP were smaller than the observed differences in the mean for SMP ratings suggesting a successful regression model.

A review of the partial regression coefficients for the 13 quality indicator ratings revealed the following four indicators as significant predictors of SMP when the impact of the other predictors was removed: Problem Validation, $t(56) = 3.05$, $p = .003$, Goal Setting $t(56) = 2.43$, $p = .018$, Intervention Plan Development, $t(56) = -2.56$, $p = .013$, and Formative Evaluation, $t(56) = 2.78$, $p = .007$. A one point increase in the Problem Validation indicator will result in a .25 ($b = .25$) increase in SMP rating when the other indicators are held constant. A one-point increase in the Goal Setting indicator will result in a .18 ($b = .18$) increase in SMP rating when the other indicators are held constant. A one point increase in Intervention Plan Development will result in a .34 decrease in SMP when the other indicators are held constant, and a one point increase in the Formative Evaluation indicator will result in a .33 ($b = .33$) increase in SMP rating when the other indicators are held constant, and. Intervention Plan Development was the only significant predictor of SMP that was also a significant predictor of GAS and was likely so for the same reason, because when teams do not report what interventions are being used they may have no real basis for determining whether or not the student made progress. The indicators Problem Validation, Goal Setting, and Formative Evaluation each fit into different behavioral consultation components. The Problem Validation indicator is part of the Problem Analysis component, while the Goal Setting indicator is part of the Plan Development component. The Formative Evaluation is part of the Plan Evaluation component. Because these indicators represent practices from along the behavioral consultation model, preintervention data collection,

intervention design, and analysis of effectiveness of postintervention implementation are all important steps when conducting GEI. Table 9 includes the partial regression coefficients for each of the 13 quality indicators.

Table 9

Summary of Multiple Regression Analysis for Variables Predicting SMP (N = 72)

Variable	*B*	*SE B*	β	*t*	*p*
Behavioral Definition	-.136	.10	-.15	-1.32	.193
Baseline Data	.030	.14	.03	0.22	.829
Problem Validation	.251	.08	.33	3.05	.003*
Problem Analysis	.028	.15	.20	0.19	.849
Goal Setting	.176	.07	.25	2.43	.018*
Intervention Plan Development	-.336	.13	-.34	-2.56	.013*
Strength of Intervention	.239	.12	.22	1.98	.052
Measurement Strategy / Progress Monitoring Plan	.140	.09	.17	1.52	.135
Decision Making Plan	-.218	.12	-.20	-1.80	.077
Progress Monitoring	-.066	.14	-.08	-0.49	.628
Formative Evaluation	.328	.12	.34	2.78	.007*
Treatment Integrity	.097	.07	.16	1.45	.152
Summative Evaluation	.181	.10	.24	1.74	.087

Note. $R^2 = .55$ for Model. Adjusted $R^2 = .45$. $df = 58$.
* $p < .05$.

Interestingly, the most important indicators significantly correlated to student performance were those that involve investigating either the problem being addressed or the intervention being implemented. None of the variables dealing with the description of one of the consultation phases was found to be significantly related to student performance. In other words, the indicators dealing specifically with objective, numerated variables were significant, while those designed to describe the student or the intervention were not.

Research Question Three

Do significant differences exist between the quality indicators of the interventions created by BBT for GEI teams and the outcomes for the three student groups? One goal of the GEI

process is to limit the number of referrals to special education, and to make sure that when referrals are made they are appropriate. One way to analyze this is to compare the process outcome variables for the students who have been in the process and were either referred for special education or exited the process because of successful intervention or other factors. Sindaler et al. (1992) recommended analyzing GEI student special education placements as a means for providing information on the overall successfulness of the GEI process.

Process Effects on Student Group

An analysis of variance (ANOVA) was conducted to compare the student groups to determine whether differences exist between the student groups on the outcome measures: students referred but not identified for special education, students referred and found eligible for special education, students no longer in the problem-solving process. Students remaining in the problem-solving process were not included in analysis because it is not known in which postprocess group these students will belong. The independent variable for the ANOVA was the student group (referred but not identified, referred and identified, or the two no longer in GEI groups). The dependent variable for the ANOVA was the overall student outcomes score obtained by summing and finding the means of the two outcome measures, Goal Attainment Scaling, and student measured performance.

Thirty-three of the 97 cases that were included in this study did not indicate the student group to which the cases were assigned. Of the 64 cases that did indicate student group, 33 indicated that the student was still in the GEI process. These cases were also excluded from the analysis because they will eventually be in one of the other student groups. This left just 31 cases for the ANOVA. The descriptives for the cases included in the ANOVA are summarized in Table 10. Due to the small number of cases in the three levels of the dependent variable, the

statistical power of the ANOVA is limited. Cohen (1992) has recommended using a statistical power of .8 or an 80% chance of detecting an effect if one exists. The statistical power of the ANOVA was .68. This statistical power means the ANOVA had less than a 70% chance of actually finding an effect if it existed and suggests the 31 cases that were included in the analysis were too few to reliably find significant results. The assumption of homogeneity of variance underlying the ANOVA was problematic. The Referred and Found Eligible group and the No Longer in Process group each had more than twice the number of cases than the Referred but Not Eligible group. There was not a significant effect of the Student Groups on the Student Outcome variable, $F(2, 28) = 1.42$, $p = .26$.

Table 10

Student Group Mean Ratings on Overall Student Outcome

Group	Mean Overall Student Outcome Ratings		
	M	*SD*	*n*
Referred but not eligible	1.80	1.04	5
Referred and found eligible	2.31	0.90	13
No longer in process	2.62	0.92	13

The ANOVA model was not statistically significant and the sample size was small, but the descriptive statistics do offer some insights. The students who were no longer in the process scored higher on the outcome variables than either of the referred student groups. This makes logical sense because students who are successful in the GEI process are likely to not get referred for special education testing. Interestingly, the students who were found eligible for special education scored slightly higher than the students who were referred but not found eligible for special education. In either case, a rating of 2 suggests that minimal, if any, gains toward the student's goal were achieved.

CHAPTER 5

DISCUSSION

The limitations of this study and the future directions are discussed. The limitations center on the type of data used (archival) and the method for rating the data (i.e., the innovation configurations). Future direction discussion notes potential future research questions and implications for BBT for GEI practices.

Summary

This study examined the effectiveness of BBT for GEI using objective, permanent product indices of intervention quality and analyzed BBT for GEI's effect on student performance outcomes and BBT for GEI's use of quality indices of intervention design. The purpose of this study was to evaluate empirically the BBT for GEI teams' effectiveness at improving student outcomes and aligning team practices with the research identified best practices. To achieve this purpose, three research questions were asked.

The first research question addressed BBT for GEI's alignment with best practices for conducting general education intervention. The question included two parts: how well do BBT for GEI teams' practices align with the quality indicators of intervention identified in the literature? What themes or commonalities are evident related to BBT for GEI teams' alignment with the quality indicators? The indicators ranged from a low rating of 1.44 to a high rating of 3.64. This range suggests that the teams implemented some of the best practice indicators to a

high degree, while other indicators were either not implemented to a high degree or not addressed. BBT for GEI teams implemented the Problem Analysis and Plan Development components with the highest fidelity while implementing the Plan Implementation and Plan Evaluation components with the lowest fidelity. When analyzing the themes and commonalities, it became apparent that many teams did not meet after their initial meeting in order to implement and monitor a plan. This partially accounted for the lower rating on the Plan Evaluation component.

The second research question was proposed to provide information on the effectiveness of BBT for GEI at improving student outcomes. The questions were: What effects do the BBT for GEI teams have on student outcomes? (a) To what degree do BBT for GEI teams' alignment with the quality indicators of intervention predict the attainment of goals set by BBT for GEI teams? (b) To what degree do BBT for GEI teams' alignment with the quality indicators of intervention predict student performance on academic or behavioral measures implemented during GEI? Each case was rated on a Likert scale from 1-5 on the case's obtainment of team stated goal. The average rating for Goal Attainment Scaling was 2.92. Ratings were assigned for Student's Measured Performance the same way as for Goal Attainment Scaling with Likert ratings from 1-5. The average rating for Student's Measured Performance ($M = 1.93$). Two multiple regression analyses were conducted to determine the effect the BBT for GEI process has on the attainment of GEI team stated goals for their students and student measured performance. The linear combination of the quality indicators of intervention design ratings was significantly related to the Goal Attainment Scaling ratings, $F(13, 58) = 2.85$, $p < .01$. The multiple correlation coefficient was .62. The percentage of variability in the Goal Attainment Scaling indicator that can be attributed to differences in Quality Indicator ratings is 39% ($R^2 = .39$). A

review of the partial regression coefficients for the 13 quality indicator ratings revealed only two indicators, Problem Analysis, $t(56) = 3.15$, $p = .003$, and Intervention Plan Development, $t(56) = -2.05$, $p = .045$, as being significant predictors of GAS when the impact of the other predictors was removed. The linear combination of the quality indicators of intervention design ratings was significantly related to the Student Measured Performance, $F(13, 63) = 5.67$, $p < .01$. The multiple correlation coefficient was .73, indicating that approximately 54% of the variance of the Measured Student Performance indicator rating can be accounted for by the linear combination of the quality indicators of intervention design. A review of the partial regression coefficients for the 13 quality indicator ratings revealed four indicators as significant predictors of SMP when the impact of the other predictors was removed: Problem Validation, $t(56) = 3.05$, $p = .003$, Goal Setting, $t(56) = 2.43$, $p = .018$, Intervention Plan Development, $t(56) = -2.56$, $p = .013$, and Formative Evaluation, $t(56) = 2.78$, $p = .007$.

The third research question was: Do significant differences exist between the quality indicators of the interventions created by BBT for GEI teams and the outcomes for the following student groups (a) students referred for a special education evaluation and subsequently identified as being eligible for special education, (b) students referred for a special education evaluation and subsequently not found eligible for special education, and (c) students no longer in GEI? An ANOVA was conducted to compare the following student groups to determine whether differences exist between the student groups on the outcome measures: students referred but not identified for special education, students referred and found eligible for special education, and students no longer in the problem-solving process due to meeting goals or due to unknown reasons. The independent variable of the ANOVA was the student group (referred but not identified, referred and identified, or the two no longer in GEI groups). The dependent

variable for the ANOVA was the overall student outcomes score obtained by summing and finding the means of the two outcome measures, goal attainment scaling, and student measured performance. Thirty-three of the 97 cases that were included in this study did not indicate the student group to which the cases were assigned. Of the 64 cases that did indicate student group, 33 indicated that the student was still in the GEI process. These cases were also excluded from the analysis because they will eventually be in one of the other student groups. This left just 31 cases for the ANOVA. Due to the small number of cases in the three levels of the dependent variable, the statistical power of the ANOVA was very limited. The assumption of normality underlying the ANOVA appeared problematic. There was not a significant effect of the Student Groups on the Student Outcome variable average for the three student groups, $F(2, 28) = 1.42$, $p = .26$.

Conclusions

The primary goal of this study was to gain information regarding the effectiveness of BBT for GEI at improving student outcomes. Previous research (Tilly & Flugum, 1995; Flugum & Reschly,1994; Telzrow et al., 2000; Upah & Tilly, 2002; & Stickler, 2004) has established that teams' outcomes are better when their practices are aligned with a behavioral consultation model including problem identification, problem analysis, plan development, and plan evaluation. Better alignment with this model was a goal of BBT for GEI. The results of this study suggest that the BBT for GEI process represents significant improvements over the CPS for GEI process due to positive alignment of BBT for GEI to the behavioral consultation model. The process aligns well with the behavioral consultation model components of problem identification and problem analysis. Improvements either in the data recording or implementation may be needed in the areas of plan development or plan evaluation. This result

is similar to other studies which found those teams' practices aligned better to problem identification and problem analysis than to the other components (Stickler, 2004; Telzrow et al., 2000).

Question two was included in the study to add to the body of research to further study how much the quality indicators contribute to student outcomes. Previous researchers have shown that greater implementation of the quality correlates to better student outcomes. Flugum and Reschly (1994) examined six quality indicators and found implementation significantly correlated with reports of improved behavior ($r = .29$) and student functioning ($r = .32$). Upah (1998) found significant positive relationship between the quality of intervention, identified through nine quality indicators, and student outcomes ranging from $r = .21$ to $r = .50$. The results of the current study found a strong relationship between the quality indicators and student outcome variables. The linear combination of the quality indicators of intervention design ratings was significantly related to the Goal Attainment Scaling ($n = 72$, $r = .62$, $p < .01$), and the linear combination of the quality indicators of intervention design ratings was significantly related to the Student Measured Performance ($n = 72$; $r = .74$, $p < .01$). These correlations cannot be compared directly to the others studies due to the different number of quality indicators and different student outcome variables, but the results do provide further support for the importance of aligning problem solving practices to the quality indicators of intervention design.

Question three was asked in order to provide additional outcome data for the BBT for GEI process. The student group data was obtained primarily to analyze whether students were appropriately referred for special education testing. This information is valuable because problem-solving teams are often used within a larger response to intervention model for

determining special education services. A goal of problem-solving teams is to reduce inappropriate referrals by providing tiers of intervention to students to remediate educational problems that are may not be so significant that they warrant special education. Problem-solving teams intervene between the tiers of response to intervention. Unfortunately, not enough teams submitted data indicating which student group their cases belonged to. The ANOVA model was not significant.

Limitations

Case Submission

The case selection process for this study was not fully randomized. The cases that were selected for inclusion in the study were taken from the group of cases submitted by participating schools. These schools undoubtedly conducted more general education interventions than what were documented in the cases submitted to the Blumberg Center. Since schools voluntarily submitted cases, no oversight was conducted to monitor which cases got submitted and which, if any, did not. It is unlikely that schools submitted only the most completed cases since many submitted cases had only partially completed the process. Knowing the extent to which the submitted cases reflect the overall cases conducted by the schools is impossible due to the voluntary submission of cases.

A second limitation related to the case submission was how students with multiple identified goals were treated in the study. Ultimately, each individually identified goal for students was treated as a unique case in the study. This was done because a way to reconcile the two goals into one rating on the innovative configuration could not be found. Separating the goals into unique cases allowed for independent ratings on each goal. Analyzing one goal per student was considered but not done because the background information related to the goals

could not be separated to include only the information related to one goal. The result of separating the multiple goals of students was creating confounding results in the analysis. Teams tended to, but not always, conduct the GEI process the same way for each goal resulting in identical ratings for the goals. This resulted in students with multiple goals receiving increased weight in the statistical analysis when their practices were identical for each goal or case and potentially resulted in a lack of independence.

Generalizability

A second limitation of this study is a reflection of its intent. That is, the results of this study are intended primarily to provide information on a specific method for conducting problem-solving intervention, BBT for GEI, and therefore any generalization beyond that can only be made with caution. More than anything, this study was designed to provide information to teams utilizing the BBT for GEI process on the extent to which they are utilizing the recognized best practices for implementation. Other teams not using the BBT for GEI process may still obtain important information from this study as to which problem solving components are most important to general education intervention. These teams should, however, determine how well their teams' activities align to the best practices first in order to provide detailed information regarding their own practices. Even teams trained in the BBT for GEI process may not be able to fully generalize these results to their practices. This is because the teams who submitted cases used in this study had only recently been trained in the process and submitted their cases from the 2007-2008 school year.

Innovation Configurations

Another limitation of this study is the nature of how the teams' adherence to the best practices for general education intervention was measured. The use of innovation configurations

provided a quantified means of rating each research-identified best practice for implementation of GEI, but the use of innovation configurations is not without limitations. First of all, the ratings of 1-5 are themselves based on subjective divisions of the indicator into five distinct levels. These levels are based on increased presence of the best practice criteria for each indicator, but the weight of importance for each criterion was not accounted for. For instance, the Intervention Plan Development indicator consisted of four criteria: (a) procedures/strategies, (b) materials, (c) when, (d) where, and (e) persons responsible. A score of 1 was given to cases that included none of these criteria. Scores on Intervention Plan Development went up by one point for each criterion that was included. The inclusion of each criterion was treated equally even though the criteria differed from one another in importance. In this example, including the procedures/strategies is likely more important when developing a plan than documenting where the intervention will take place, but these two criteria are weighed equally when rating the innovation configuration for the Intervention Plan Development indicator.

An additional limitation of the innovation configurations is evident when attempting to compare ratings to one another. Each innovation configuration was constructed to measure the adherence to the best practices for each quality indicator of intervention design. As such, teams' practices with one indicator can not necessarily be compared to their practices with other indicators. For example, an average rating of 3.5 on the Baseline Data indicator and a 2.5 on the Problem Analysis can only be interpreted as teams' practices on Baseline Data are more closely aligned to best practices for baseline data than their practices on Problem Analysis are aligned to problem analysis practices.

Documentation of Activities

The data (e.g., the rating forms and documents) submitted by the teams to the Blumberg Center also had limitations. The data only represents what each team documented during the GEI process. Teams may not have documented their practices completely or accurately. In fact, not all the teams' practices could be documented on the forms submitted to the Blumberg Center. Team deliberations, standardized testing data, and teacher behavior reports are a few of the activities teams engage in that may not be documented on the forms submitted to the Blumberg Center.

Lack of Independence

Each of the 13 indicators of quality intervention design was assigned to one of the five problem solving component they matched. This could be done because the indicators overlapped procedures. This overlap meant that the indicators were related to each other. For example, the indicators Behavioral Definition and Baseline Data are both part of the Problem Identification component of problem solving. Teams that defined the problems in behavioral terms were given a foundation for collecting baseline data that was objective and quantifiable. The indicators that measured team practices in plan evaluation were also related. Teams that did not meet for follow-up to discuss the implemented plans received low ratings on each of the three indicators that made up the Plan Evaluation component, Formative Evaluation, Treatment Integrity, and Plan Evaluation. The reason a lack of independence for the indicators is considered a limitation is due to the fact that the GEI process is a stepwise process and each procedure builds on the procedures before. One component or quality indicator cannot be analyzed independent of the others.

Future Directions

Future Research

This study was primarily designed to evaluate the alignment of the BBT for GEI process to the research-identified best practices for conducting problem solving teams. In the future, this design may be adapted to a more experimental design. One limitation of this study was the amount of unknown variables related to team activities because the researcher did not collect data as the teams were deliberating and conducting the GEI process. In the future, a researcher may take the BBT for GEI Action Plan Evaluation Worksheet into GEI meetings and base his or her ratings on the quality indicators on observational data as well as archival data. This would allow the researcher remove two of the limitations noted earlier, case submission and documentation of activities.

The Goal Attainment Scaling and Student Measured Performance data related to which indicators of intervention design contributed most to student outcomes. This data raised various questions that may be explored in the future. Why were some variables more related to student outcomes than others? Would these results hold true if the research was replicated? Can those variables that did not significantly predict student outcomes be removed from the GEI process altogether? The body of research, not just this study, has only begun to empirically evaluate problem solving teams' effect on student outcomes. The method of implementing and the goal of problem-solving teams are quickly shifting to fit within the Response-to-Intervention model and research on these teams' effectiveness must keep up to provide practitioners needed information regarding what works and what does not. Teams have only a finite amount of time to conduct problem solving practices and they need to know that their practices are up to date and empirically supported.

This study only evaluated the BBT for GEI process, but how would the BBT for GEI Action Plan Evaluation Worksheets be scored on other processes for conducting GEI? How would other processes' alignment, and BBT for GEI Action Plan Evaluation Worksheet ratings, compare to those of the BBT for GEI process? These questions are relevant to the fact that the BBT for GEI process has not been extensively researched, especially since its overhaul to align to a problem solving model has happened. Other student outcome research such as its impact on referral rates, may be beneficial for further evaluating BBT for GEI's overall effectiveness.

Implications for BBT for GEI practice

This study has outlined how the GEI process matches up with what are considered best practices for implementing problem solving teams. Teams that conduct BBT for GEI perform similarly to other problem solving processes in terms of this alignment. When the BBT for GEI process was created, a focus of the process was on providing data driven decision making. This focus appears to have paid off. Teams received average ratings of better than 3.0 on the quality indicators of Baseline Data, Problem Validation, Intervention Plan Development, Problem Analysis, and Goal Setting. To receive average ratings of higher than 3.0, means that BBT for GEI teams are including multiple best practice components in their practices. A focus moving forward must be on progress monitoring and plan evaluation. Teams included noticeably fewer best practice components for those indicators that make up activities related to plan evaluation and progress monitoring. The quality indicators that make up the plan evaluation and progress monitoring components are Progress Monitoring, Formative Evaluation, Treatment Integrity, and Summative Evaluation. The implication of this is that if teams do not reflect on the plan they have implemented, then they are not able to make informed decisions on the plan's effectiveness and subsequent student interventions.

REFERENCES

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

Baer, D., Wolf, M., & Risley, T. (1968). Some current dimensions of applied behavioral analysis. *Journal of Applied Behavioral Analysis, 1,* 91-97. doi: 10.1901/jaba.1987.20-313.

Bahr, M. W., Walker, K., Hampton, E. M., Buddle, B. S., Freeman, T., Ruschman, N., … Littlejohn, W. (2006). Creative problem solving for general education intervention teams. *Remedial and Special Education*, *27*(1), 27-41. Retrieved from http://sed.sagepub.com/

Bahr, M. W., Whitten, E., Dieker, L., Kocarek, C. E., & Manson, D. (1999). A comparison of school-based intervention teams: Implications for educational and legal reform. *Exceptional Children*, *66*, 67-83. Retrieved from http://www.cec.sped.org/

Baker, S., Gersten, R., & Grossen, B. (2002). Interventions for students with reading comprehension problems. In M. Shinn, H. Walker, & G. Stoner (Eds.), *Interventions for academic and behavior problems II: Preventive and remedial approaches*. (pp. 731-755). Bethesda, MD: National Association of School Psychologists.

Bandura, A. (1986). *Social foundations of thought and actions: A social cognitive theory.* Englewood Cliffs, NJ: Prentice Hall.

Barnett, D. (2002). Best practices in collaboratie problem solving for instructional design. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology – IV* (pp. 247-262). Washington, DC: National Association of School Psychologists.

Batsche, G. M., & Knoff, H. M. (1995). Best practices in linking assessment to intervention. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology – III* (pp. 569-585). Washington, DC: National Association of School Psychologists.

Bergan, J. R. (1977). *Behavioral consultation*. Columbus, OH: Charles E. Merrill.

Bergan, J. R., & Kratochwill, T. R. (1990). *Behavioral consultation and therapy*. New York, NY: Plenum.

Blumberg Center for Interdisciplinary Studies in Special Education. (2006). *BBT for GEI training module.* Unpublished Manuscript, Blumberg Center for Interdisciplinary Studies in Special Education, Indiana State University, Terre Haute, IN

Buck, G., Polloway, E., Smith-Thomas, A., & Cook, K. (2003). Prereferral intervention processes: A survey of state practices. *Exceptional Children, 69,* 349-351. Retrieved from http://www.cec.sped.org/

Burns, M. K., Vanderwood, M. L., & Ruby, S. (2005). Evaluating the readiness of pre-referral intervention teams for use in a problem solving model. *School Psychology Quarterly, 20*, 88-105. doi: 10.1521/scpq.20.1.89.64192

Burry-Stock, J. A., Shaw, P. G., Laurie, C., & Chissom, B. (1996). Rater agreement indexes for performance assessment. *Educational and Psychological Measurement, 56,* 251-262. doi: 10.1177/0013164496056002006

Cantrell, R. P., & Cantrell, M. L. (1980) Ecological problem solving: A decision-making heuristic for prevention-intervention education strategies. In J. Hogg & P. J. Mittler (Eds.), *Advances in mental handicap research* (pp. 267-301). New York, NY: Wiley.

Carter, J., & Sugai, G. (1989). Survey on prereferral practices: Responses from state departments of education. *Exceptional Children*, *55*, 298-302. Retrieved from http://www.cec.sped.org/

Chalfant, J. C., & Pysh, V. (1989). Teacher assistance teams: five descriptive studies on 96 teams. *Remedial and Special Education, 10*, 49-58. doi: 10.1177/074193258901000608

Cobb, C. T., (1995). Best practices in linking assessment to intervention. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology–III* (pp. 325-336). Washington, DC: National Association of School Psychologists.

Cohen, J., (1992). A power primer. *Psychological Bulletin*, *112*, 155-159. doi: 10.1037/0033-2909.112.1.155

Deno, S. (1995). Best practices in facilitating intervention adherence. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology–III* (pp. 471-484). Washington, DC: National Association of School Psychologists.

Education for All Handicapped Children Act, 20 U.S.C. §1400 *et seq*.(1975).

Flugum, K. R., & Reschly, D. J. (1994). Prereferral interventions: Quality indices and outcomes. *Journal of School Psychology, 32*, 1-14.

Fuchs, D., & Fuchs, L. (1989). Exploring effective and efficient prereferral interventions: A component analysis of behavioral consultation. *School Psychology Review*, *18*, 260-279.

Fuchs, D., Fuchs, L., & Bahr, M. (1990). Mainstream assistance teams: a scientific basis for the art of consultation. *Exceptional Children, 57*(1)*,* 128-139. Retrieved from http://www.cec.sped.org/

Fuchs, D., Fuchs, L., Gilman, S., Reeder, P., Bahr, M., Fernstorm, P, & Roberts, H. (1990). Prereferral intervention through teacher consultation: mainstream assistance teams. *Academic Therapy*, *25*, 263-276. Retrieved from http://www.academictherapy.com/

Gettinger, M., & Stoiber, K. (1999). Excellence in teaching: Review of instructional and environmental variables. In C. R. Reynolds & T. B. Gutkin (Eds.), *The handbook of school psychology-3rd Edition* (pp. 933-958). New York, NY: John Wiley.

Graden, J. L., Casey, A., & Christenson, O. (1985). Implementing a prereferral intervention system: part I the model. *Exceptional Children*, *51*, 377-384. Retrieved from http://www.cec.sped.org/

Gravois, T. A., & Rosenfield, S. (2002). A multidimensional framework for evaluation of instructional consultation teams. *Journal of Applied School Psychology*, *19*(1), 5-29. Retrieved from http://www.taylorandfrancisgroup.com/

Gresham, F. M. (2001). Responsiveness to intervention: An alternative approach to the identification of learning disabilities. In. R. Bradley, L Danielson, & D. Hallahan (Eds.), *Identification of learning disabilities: Research to practice, 6*. Mahwah, NJ: Lawrence Erlbaum.

Hampton, E. M. (2004). CPS for GEI district and school-based team focus group report executive summary. Retrieved from http://www.indstate.edu/blumberg/presm/publications-briefs.htm

Hord, S. M., Rutherford, W., Huling-Austin, L., & Hall, G. (1987). *Taking charge of change.* Austin, TX: Southwest Education Development Laboratory.

Horn, L. J., & Carroll, C. D. (1997). *Confronting the odds: Students at-risk and in the pipeline to higher education*. (Report No. 98-094) Washington, DC: National Center for Education Statistics. Retrieved from National Center for Education Statistics website: http://nces.ed.gov/pubs98/98094.pdf

Howell, K. W., Kurns, S., & Antil, L. (2002). Best practices in problem-solving team structure and process. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology – IV* (pp. 753-788). Washington, DC: National Association of School Psychologists.

Indiana State Board of Education: Special Education Rules, 511 IAC § 7-17-40 (2002).

Individuals with Disabilities Education Act, 20 U.S.C. § 1400 *et seq.* (1990).

Individuals with Disabilities Education Act Amendments of 1997, 20 U.S.C. § 1400 *et seq.* (1997).

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

Ivarie, J., & Russell, J. (1992, November). *School-based teams: Can they increase "verifiable" referrals to special education.* Paper presented at the meeting of the Teacher Education Division of the Council for Exceptional Children Conference, Cincinnati, OH. Retrieved from http://www.cec.sped.org/

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. 657-670). Washington, DC: National Association of School Psychologists.

Kazdin, A. E. (1982). *Single-case research designs: Methods for clinical and applied settings.* New York, NY: Oxford University Press.

Kovaleski, J. F. (2002). Best practices in operating problem-solving intervention teams. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology – IV* (pp. 645-655). Washington, DC: National Association of School Psychologists.

Kratochwill, K., Elliott, G., & Callan-Stoiber, K. (2002). Best practices in school-based problem-solving consultation. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology – IV* (pp. 583-608). Washington, DC: National Association of School Psychologists.

Kruger, L. J., Struzziero, J., Watts, R., & Vacca, D. (1995). The relationship between organizational support and satisfaction with teacher assistance teams. *Remedial and Special Education*, *16*, 203-211. doi: 10.1177/074193259501600404

Lane, K. L., Pierson, M. R., Robertson, E. J., & Little, A. (2004). Teacher's views of prereferral interventions: perceptions of and recommendations for implementation support. *Education and Treatment of Children*, *27*, 420-439. Retrieved from http://www.educationandtreatmentofchildren.net

Macan, T. H. (1994). Time management: A test of a process model. *Journal of Applied Psychology*, *79*, 381-391. doi: 10.1037/0021-9010.79.3.381

Menard, S. (1995). *Applied logistic regression analysis*. Thousand Oaks, CA: Sage.

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

Ross, R. P. (1995). Best practices in linking assessment to intervention. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology – III* (pp. 227-237). Washington, DC: National Association of School Psychologists.

Shinn, M. R. (2002). Best practices in using curriculum-based measurement in a problem-solving model. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology – IV* (pp. 671-697). Washington, DC: National Association of School Psychologists.

Sindaler, P. T., Griffen, C. C., Smith, S. W., & Wantanabe, A. K. (1992). Prereferral intervention: encouraging notes on preliminary findings. *The Elementary School Journal*, *92*, 245-259. doi: 10.1086/461691

Slonski-Fowler, K. E., & Truscott, S. D. (2004). General education teachers' perceptions of the prereferral intervention team process. *Journal of Educational and Psychological Consultation*, *15*, 1-40. Retrieved from http://www.leaonline.com

Snow, C., Burns, S., & Griffin, P. (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

Steege, M., Brown-Chidsey, R., & Mace, F. C. (2002). Best practices in evaluating interventions. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology-IV* (pp. 517-534). Washington, DC: National Association of School Psychologists.

Strickler, W. L. (2004). *A descriptive study of indicators of effective intervention for supporting students* (Unpublished doctoral dissertation). University of Cincinnati, Cincinnati, OH.

Stokes, T. F., & Baer, D. (1977). An implicit technology of generalization. *Journal of Applied Behavioral Analysis, 10, 349-367.* doi: 10.1901/jaba.1977.10-349

Telzrow, C. F. (1995). Best practices in facilitating intervention adherence. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology–III* (pp. 501-518). Washington, DC: National Association of School Psychologists.

Telzrow, C. F., McNamara, K., & Hollinger, C. L. (2000). Fidelity of problem-solving implementation and relationship to student performance. *School Psychology Review, 29,* 443-461.

Tilly, W. D., III, & Flugum, K. R. (1995). Best practices in ensuring quality interventions. In A. Thomas & J. Grimes (Eds.), *Best Practices in school psychology – III* (pp. 485-500). Washington, DC: National Association of School Psychologists.

Truscott, S. D., Cohen, C. E., Sams, D. P., Sanborn, K. J., & Frank, A. J. (2005). The current state(s) of prereferral intervention teams: A report from two national surveys. *Remedial and Special Education*, *3*(1), 130-140. doi: 10.1177/07419325050260030101

United States Department of Education. (1998). *Twentieth annual report to congress on implementation of the individuals with disability education act.* Washington, DC: Author.

Upah, K. R F. (1998). *School-based problem-solving interventions: The impact of training and documentation on the quality and outcomes* (Unpublished doctoral dissertation). Iowa State University, Ames, IA.

Upah, K. R. F., & Tilly, W. D., III, (2002). Best practices in designing, implementing, and evaluating quality interventions. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology – IV* (pp. 483-499). Washington, DC: National Association of School Psychologists.

Vaughn, S., Hughes, M., Schumm, J. S., & Klinger, J. (1998). A collaborative effort to enhance reading and writing instruction in inclusion classrooms. *Learning Disability Quarterly, 21*(1), 57-74. Retrieved from http://www.cldinternational.org/Publications/LDQ.asp

Yetter, G., & Doll, B. (2007). The impact of logistical resources on prereferral team acceptability. *School Psychologist Quarterly, 22*, 340-357. doi: 10.1037/1045-3830.22.3.340

Zins, J. E., & Erchul, W. P. (2002). Best practices in school consultation. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology – IV* (pp. 625-644). Washington, DC: National Association of School Psychologists.

APPENDIX A: BBT FOR GEI EVALUATION WORKSHEET

BBT for GEI Action Plan Evaluation Worksheet

For each case file, complete one worksheet for the primary target behavior. Complete only one worksheet per case file, unless it is not possible to determine which of multiple behaviors is the primary or priority behavior carried throughout the intervention. In this case, complete one worksheet for each target behavior that is carried throughout problem solving (i.e. has a behavioral definition, goal, intervention plan, and individual graph).

For case files containing multiple levels/trials of interventions, complete the worksheet using the problem-solving form and set of interventions developed only through documented problem solving (not including interventions conducted prior to problem solving).

Case ID number: _____ School Name: _____

Rater initials: _____ Date worksheet completed: _____

Student: Grade: _____ Race/Ethnicity: _____ Gender: __

Reason for Referral: _____
_____

Concern (Check primary concern):
____: Reading
____: Written Expression
____: Oral Expression
____: Listening Comprehension
____: Math
____: Other_____
____: Learning Behaviors
____: Social Behaviors

Number of intervention trials/levels completed: _____
(# of Intervention/Action Plan forms from different dates, indicating reviews/changes to the plan)

Check all of the following products found in the case file:
____: Background Information/Referral Form
____: Intervention/Action Plans
____: Team Accomplishment Sheet

Directions for Rating:

- For each item with an innovation configuration component, review all evidence in the case files and record the degree of implementation based upon the predetermined categories in the innovation configuration.
- Refer to asterisks for each innovation configuration for where to find data for ratings
- After rating each innovation configuration, in the accompanying box, write a narrative explaining why each configuration received the rating it did.
- Upon completion of the 12 items with an innovation configuration component, complete the case summary table

Quality Indicators of Intervention Worksheet

1.) Behavioral Definition*

5	4	3	2	1
Definition is (a) objective—refers to observable and Measurable characteristics of behavior, (b) clear—written so that it is easily understood, repeated, and paraphrased by observers (c) complete—delineates both examples and non-examples of the behavior.	Definition meets only two of the three criteria (i.e. objective, clear, and complete)	Definition meets only one of the three criteria (i.e. objective, clear, and complete)	Problem behavior is stated in general terms (e.g. reading comprehension, aggressive behavior, etc.).	Behavioral definition is not written.

*refer to Goal and Critical questions on Action Plan for ratings

1.) Behavior Definition (narrative)

2.) Baseline Data*

5 Data collected on the behavior prior to implementing the intervention consisting of repeated measures of the behavior over several (at least three) sessions, days, or even weeks until a stable range of behavior has Exceptions: 1-2 data points or reconstructed baseline for dangerous behavior	4 Data collected on the behavior prior to implementing the intervention and is written in qualitative form; however, only two data points are reported or single data points from multiple measures.	3 Data collected on the behavior prior to implementing the intervention; however, only one data point is reported or data is presented in quantitative form.	2 Data collected on the behavior prior to implementing the intervention; however, the dimensions(s) addressed are not the most appropriate for the selected behavior.	1 Baseline data not gathered prior to implementing the intervention.

* refer to Action Plan or Team Accomplishment Sheet for ratings

2.) Baseline Data (narrative)

3.) Problem Validation*

5 The magnitude of the discrepancy is quantified, based on a comparison between student's performance and the local educational benchmarks or national/out-of district benchmarks.	4 The magnitude of the discrepancy is quantified, based on a comparison between the student's performance and standards; however, it is not clear how the standard was determined (peers? Local benchmark? National benchmark?).	3 The magnitude of the discrepancy is quantified, but is based on an opinion.	2 The magnitude of the discrepancy is described qualitatively.	1 Problem is not validated; magnitude of the discrepancy is not described.

*refer to Background Information/Referral Form and Baseline Data for ratings

3. Problem Validation (narrative)

4.) Problem Analysis*

5 Assessment is linked to the target behavior; examined relevant and alterable factors in at least three of the following domains: (a) instructional access, (b) content variables, (c) environment, and (d) personal**; using a variety of procedures/ sources; and includes a functional hypothesis of the problem behavior.	4 Assessment is linked to the target behavior; examined relevant and alterable factors in at least three of the following domains: (a) instructional access, (b) content variables, (c) environment, and (d) personal; using a variety of procedures/ sources	3 Assessment is linked to the target behavior; and examined relevant and alterable factors from only two of the domains using at least two sources	2 Assessment is linked to the target behavior; examined relevant and alterable factors from only one of the domains; or only a single source was used to examine relevant factors	1 Problem analysis is not conducted or assessment was not directly linked to target behavior.

** Domain definitions:
Instructional Access: Evidence that previous instruction/interventions were used to address/remediate target behavior
Content Variables: Evidence assessment was conducted to differentiate target behavior from other behaviors or to provide more in-depth information on behavior
Environment: Evidence that environmental factors were considered (i.e., classroom environment, attendance records, ect.)
Personal: Evidence that personal factors were considered (i.e., home environment, cultural factors, health history, ect.).

4.) Problem Analysis (narrative)

5.) Goal Setting*

5 Goal specifying time frame, condition, behavior, and criteria which are based on a comparison between the student's baseline data and the expectations/ benchmarks.	4 Goal specifies three of the four components (time frame, behavior, criterion, and condition), or includes all four components but does not link baseline data to goal	3 Goal specifies two of the four components (time frame, behavior, criterion, and condition)	2 Goal specifies one of the four components (time frame, behavior, criterion, and condition)	1 Goal is not measurable or is not set.

*refer to Action Plan for ratings

5.) Goal Setting (narrative)

6.) Intervention Plan Development*

5 Plan stated procedures/ strategies, or is implied if standard procedures for delivery are known and all of the following components are included (a) procedures/ strategies, (b) materials, (c) when, (d) where, and (e) persons responsible	4 Plan stated procedures/ strategies, or is implied if standard procedures for delivery are known. But one of the following components is missing: materials, when, where, or person responsible.	3 Plan stated procedures/ strategies, or is implied if standard procedures for delivery are known. But two of the following components is missing: materials, when, where, or person responsible.	2 Generic description of intervention strategy (e.g. behavior mgmt.) is stated. Materials, when where, person responsible may be present.	1 Intervention plan not written. Or, generic description s of interventions (e.g., behavior management) only

*refer to Action Plan for ratings

6.) Intervention Plan Development (narrative)

7.) Strength of Intervention*

5	4	3	2	1
Three or more elements of strong intervention** included in the Action Plan.	At least two elements of strong intervention included in the Action Plan.	At least one element of strong intervention included in the Action Plan.	No elements of Strong interventions, but a plan that an individual has tried previously with success.	No elements of strong interventions included in the Action Plan.

*refer to Action Plan for ratings

**Elements of strong interventions:

<u>Academic</u>
Direct instruction
Opportunities to practice
Modeling/prompting
Feedback (positive contingencies)
Feedback (error correction)
Naturalistic

<u>Behavior</u>
Direct instruction
Opportunities to practice
Modeling/prompting
Feedback (contingent reinforcement)
Feedback (contingent consequence)
Naturalistic

7.) Strength of Intervention (narrative)

8.) Measurement Strategy / Progress Monitoring*

5	4	3	2	1
A measurement strategy is developed answering How? What? Where? Who? When?	A measurement strategy is developed, but only answers four of the five: How? What? Where? Who? When?	A measurement strategy is developed, but only answers three of the five: How? What? Where? Who? When?	A measurement strategy is developed, but only answers two of the five: How? What? Where? Who? When?	Measurement strategy is not developed. Or the measurement strategy only answers one of the five questions.

*refer to Action Plan for ratings

8.) Measurement Strategy / Progress Monitoring (narrative)

9.) Decision-making Plan*

5 The Action Plan indicates (a) how frequently data will be collected, (b) the strategies to be used to summarize the data for evaluation, (c) how many data points or how much time will occur before data will be analyzed, and (d) what actions will be taken based on the intervention data	4 The decision-making plan indicates three of the four components	3 The decision-making plan indicates two of the four components	2 The decision-making plan indicates only one of the four components	1 The decision-making plan is not documented.

*refer to Action Plan for ratings

9.) Decision-making Plan (narrative)

10.) Progress Monitoring*

5	4	3	2	1
Data is collected consistently, at least once per week, using the same measurement tool and is charted or graphed using appropriately conventions**	Data is collected consistently using the same measurement tool, but is not charted or graphed using appropriate conventions	Data is collected multiple times, but is either inconsistently measured or measured using different measurement tools	Progress monitoring data consisted of only preintervention and postintervention data	Progress monitoring data was not collected

*refer to Action Plan and any supplied graphed information for ratings
**Charting/Graphing Conventions: Descriptive title, Understandable scale captions, Appropriate scale units, Intervention and unplanned phase change lines
Phases labeled, Phase lines indicating intervention changes or unplanned changes (e.g. new teacher), Key if there are multiple data sets

10.) Progress Monitoring (narrative)

11.) Formative Evaluation*

5 There is evidence the decision rule was followed. The data were used to modify or change the intervention as necessary	4 There is evidence the decision rule was followed, but the data were not used to modify or change the intervention as necessary, or no decision rule was included but team made changes based on data.	3 Modifications or changes were made to the intervention based on subjective data	2 Modifications or changes were made to the intervention, but no indication as to what data were used to make these changes.	1 No formative evaluation was conducted

*refer to Action Plan and Team Accomplishment Sheet for ratings

11.) Formative Evaluation (narrative)

12.) Treatment Integrity

5 Treatment Integrity plan is noted with an implementation at or above 76%	4 Treatment Integrity plan is noted with an implementation rate between 51% and 75%	3 Treatment Integrity plan is noted with an implementation rate between 26% and 50%	2 Treatment Integrity plan is noted with an implementation rate of 25% or less	1 Treatment integrity is not considered

12.) Treatment Integrity (narrative)

13.) Summative Evaluation*

5 Outcome decisions are based on the progress monitoring data.	4 Outcome decisions are based on minimal data (i.e., pre and posttests).	3 Outcome decisions are based on subjective data.	2 Outcome decision stated but no indication of what data were used to make the conclusion.	1 No summative evaluation took place.

*refer to Action Plan and Team Accomplishment Sheet for ratings

13.) Summative Evaluation (narrative)

Case Summary Table

Indicator	Innovative Configuration Rating
1.) Behavioral Definition	
2.) Baseline Data	
Problem Identification Score	Mean Score:
3.) Problem Validation	
4.) Problem Analysis	
Problem Analysis Score	Mean Score:
5.) Goal Setting	
6.) Intervention Plan Development	
7.) Strength of Intervention	
Plan Development Score	Mean Score:
8.) Measurement Strategy	
9.) Decision-Making Plan	
10.)Progress Monitoring	
Plan Implementation Score	Mean Score:
11.) Formative Evaluation	
12.) Treatment Integrity	
13.) Summative Evaluation	
Plan Evaluation Score	
OVERAL QUALITY SCORE	Mean Score:

BBT for GEI Student Outcome Data Worksheet

Is the student still in GEI? Y N

Was a referral for Special Education testing made? Y N

If a referral was made:
Has the student been found to be eligible for special education? Y N Pending

1. Goal Attainment Scaling

5	4	3	2	1
Set goal was met or exceeded in both time frame and criterion	Set criterion goal was met or exceeded, but not within the time frame established. (If goal was achieved but no time frame was noted in goal, score = 4)	Data indicate that progress was made toward goal, but set goal was not met.	Data indicate that level of progress did not change from baseline data level.	Data indicate that level of progress decreased from baseline data level.

1. Goal Attainment Scaling (narrative)

2. Student Measured Performance

5	4	3	2	1
A change in magnitude and rate was evident across all of the following visual analysis criteria*: a) mean, b) level, c) trend, and d) latency	A change in magnitude and rate was evident across 3 of the four visual analysis criteria*	A change in magnitude and rate was evident across 2 of the four visual analysis criteria*	A change in magnitude and rate was evident for only 1 of the four visual analysis criteria*	A change in magnitude and rate was not evident in any of the four visual analysis criteria *

* Visual analysis criteria:
Mean –the average performance rating is higher during the intervention phase than the baseline data phase
Level –there is discontinuity of performance (in the desired direction) from the end of the baseline to the start of the intervention
Trend –the trend in performance increase over time
Latency –there a change in performance (in the desired direction) after the first week of intervention implementation

2. Changes in magnitude and rate (narrative)

Intervention Implementation Fidelity Checklist

Student: Date:

	WK	WK2	Wk3	Wk4	WK5	WK6	WK7	WK8	WK9	Summary
Intervention										
a. Implemented as Defined										
b. Frequency										
c. Duration										
d. Interventionist (Who?)										
e. PM Data collected as Defined										
Completed By: Date: Notes:										

APPENDIX C: INTERATER RELIABILITY AID

Considerations for Making Ratings on the Intervention Worksheet

Item	Place to obtain rating	Additional scoring considerations/reminders
1.	Use action plan "Area of Concern/ Subskill" first to obtain rating. If this is left blank, use "Reason for Referral" on Background Information/Referral Form for rating.	
2.	Use action plan "Baseline Data" first to obtain rating. Assessment data in the Background Information Form may be considered baseline data only if it is referenced in the Action Plan as such.	Data must be presented quantitatively to earn scores of 3 or higher. A score of 5 must use repeated measures of same assessment (more than three). A score of 4 must have multiple measures or two data points from a single measure.
3.	Use Background Information/ Referral Form for rating.	The magnitude of the discrepancy must be between the identified problem behavior and some standard. Qualitative magnitude includes information such as failing grades, Likert ratings, and statements of severity.
4.	Use Background Information/	A score of five must include a written functional

	Referral Form for rating and baseline data. Also consider critical questions.	hypothesis of the target behavior. There must be evidence that the assessment information was used to rule in/out possible factors for the target behavior to obtain a rating of greater than 1. If only one assessment source was used to analyze the problem, the highest rating that may be given is 2. When analyzing assessment linkage, consider these examples of the alterable factors (variables): curriculum – student was previously taught the skill (was student exposed to materials). Instruction – team lists previous instructional strategies to teach skill to student (teacher tried different methods of getting materials across). Environment – classroom behaviors are considered as alternative functions of behavior. Student – physical limitations, acculturation, home life variables.
5.	Action plan for ratings.	Goal condition: circumstance under which the goal will be measured is evident (e.g., to improve score on a specific test or reduce behavior in a specific setting). Goal criteria: observable/objective point when goal is met is established and evident. Some teams do not re-list all the goal components on the goal line in the action plan, but should be given credit for including them if they are included elsewhere (e.g., teams often will write in the date the team is going

		to meet and review progress. This may be considered the time frame even though it isn't included in the goal statement).
6	Action plan	If team lists a packaged intervention, or one that clearly implies standardized procedures (e.g., math computer program, or morning classroom diary), do not count as a generic description only. Many teams list multiple interventions or also list accommodations. Do not base ratings on listed accommodations. For multiple interventions for a single goal, count intervention components in each listed intervention and base rating on the team's overall alignment to the quality indicator.
7	Action plan	If there are multiple interventions listed for the goal, count the different strong elements from all of the interventions to obtain the rating (i.e., if there are three interventions listed and those interventions included 2 different strong elements give the team a 4. Note that repeated use of a strong intervention element is only counted as one.
8.	Action plan	Consider the following for obtaining ratings: How? = procedures for measuring. What? = what measurement is being used (does not have to be detailed if standard procedures are implied or referenced… DIBELS, grades, attendance, etc.) Where? = where will the

		measurement take place. Who? = Who is responsible for obtaining the measurements. When? = when will measurements be taken.
9	Action plan and goal statement.	
10.	Action plan, implementation fidelity summary data, team accomplishment sheet	
11.	Action plan, team accomplishment sheet.	Formative evaluation takes progress monitoring a step further, answering the question 'Is the plan working?' and using this outcome information to adjust plans during the implementation phase
12	Action plan, implementation fidelity summary data sheet. Implementation fidelity summary data column.	
13.	Action plan. Team accomplishment sheet	Summative evaluation refers to a comparison of baseline and postintervention performance. This is an examination of the data after the planned interventions, to determine whether the intervention was as effective as intended.
Goal attainment scaling	Action plan, team accomplishment sheet	If there was no goal stated, rating cannot be higher than a 2.

Student measured performance	Action plan, team accomplishment sheet	The teams do not need to report their data visually. If team presents pre- and postimprovement only, than score may be a 2.