A meta-analysis of functional behavioral assessments of students with high-incidence disabilities in public education settings

Jeff Gabrielson

Follow this and additional works at: https://repository.usfca.edu/diss

Part of the Education Commons

Recommended Citation
https://repository.usfca.edu/diss/250
A META-ANALYSIS OF FUNCTIONAL BEHAVIORAL ASSESSMENTS OF
STUDENTS WITH HIGH-INCIDENCE DISABILITIES
IN PUBLIC EDUCATIONAL SETTINGS

A Dissertation
Presented to the Faculty of the School of Education
Learning and Instruction Department

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Education

by
Jeff Gabrielson
San Francisco
May 2007
This dissertation, written under the direction of the candidate's dissertation committee and approved by the members of the committee, has been presented to and accepted by the Faculty of the School of Education in partial fulfillment of the requirements for the degree of Doctor of Education. The content and research methodologies presented in this work represent the work of the candidate alone.

Jeff Gabrielson  
Candidate

May 7th, 2007  
Date

Dissertation Committee

Patricia Buck  
Chairperson

7 May 2007

Susan Evans

5.7.07

Steven Hartung

5.7.07
I would like to take this opportunity to acknowledge a few individuals who have supported me in the completion of this thesis. First, I would like to thank my wife Margarett for her unwavering commitment in allowing me to achieve my personal and professional goals while providing a supportive and nurturing environment for our family. Through this process, she has endured countless weekend and weeknight classes, given birth to two of our children, made two geographic moves, and embarked upon her own professional journey—now it is your turn. Second, I would like to acknowledge my wonderful children Raven, Azul, Georgia, and Ryder for whom my only wish is that they pursue their dreams, whatever they may be, and that they know how proud I am to be their father. Third, I want to express my gratitude to my father, Rod, for his support and guidance through the years. Thank you for being such a great role model, I will always admire your strong sense of ethics and sincerity. Lastly, I would like to thank my advisor and committee chairperson Dr. Patricia Busk whose incredible flexibility, consistency, and guidance has made the completion of this thesis possible. Thank you for assisting me in this process, I have learned so much.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TABLE OF CONTENTS</td>
<td>iii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>vi</td>
</tr>
<tr>
<td>List of Figures</td>
<td>vii</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td>1</td>
</tr>
<tr>
<td>Background and Need</td>
<td>6</td>
</tr>
<tr>
<td>The Context of Educating Students with Special Needs</td>
<td>7</td>
</tr>
<tr>
<td>Functional Behavioral Assessment</td>
<td>11</td>
</tr>
<tr>
<td>The Evolution of FBA</td>
<td>15</td>
</tr>
<tr>
<td>Emergence of the Multiple Methods of FBA</td>
<td>18</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td>29</td>
</tr>
<tr>
<td>Research Questions</td>
<td>30</td>
</tr>
<tr>
<td>Significance of the Study</td>
<td>30</td>
</tr>
<tr>
<td>Theoretical Rationale</td>
<td>31</td>
</tr>
<tr>
<td>Definition of Terms</td>
<td>35</td>
</tr>
<tr>
<td>Summary</td>
<td>39</td>
</tr>
<tr>
<td>CHAPTER</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>II. REVIEW OF THE LITERATURE</td>
<td>41</td>
</tr>
<tr>
<td>Effectiveness of PBS and FBA</td>
<td>41</td>
</tr>
<tr>
<td>Synthesis, Descriptive, and Meta-analytic Studies Related to and FBA methods the Efficacy of Behavioral Interventions</td>
<td>56</td>
</tr>
<tr>
<td>Interventions to decrease problematic behavior in schools</td>
<td>58</td>
</tr>
<tr>
<td>FBA methods</td>
<td>67</td>
</tr>
<tr>
<td>Descriptive analysis and critique of literature on school-based FBAs</td>
<td>71</td>
</tr>
<tr>
<td>The utility, acceptability, and practicality of FBA</td>
<td>75</td>
</tr>
<tr>
<td>Summary</td>
<td>79</td>
</tr>
<tr>
<td>III. METHODOLOGY</td>
<td>83</td>
</tr>
<tr>
<td>Method</td>
<td>83</td>
</tr>
<tr>
<td>Strategies for Literature Search</td>
<td>85</td>
</tr>
<tr>
<td>Inclusion Criteria</td>
<td>86</td>
</tr>
<tr>
<td>Variables</td>
<td>88</td>
</tr>
<tr>
<td>Effect-size Measure</td>
<td>88</td>
</tr>
<tr>
<td>Coding Sheet</td>
<td>89</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>97</td>
</tr>
<tr>
<td>IV. RESULTS</td>
<td>101</td>
</tr>
<tr>
<td>Research Question 1</td>
<td>107</td>
</tr>
<tr>
<td>Research Question 2</td>
<td>109</td>
</tr>
</tbody>
</table>
**TABLE OF CONTENTS CONTINUED**

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research Question 3</td>
<td>112</td>
</tr>
<tr>
<td>Research Question 4</td>
<td>113</td>
</tr>
<tr>
<td>Summary</td>
<td>116</td>
</tr>
<tr>
<td>V. DISCUSSION, RECOMMENDATIONS, AND CONCLUSIONS</td>
<td>118</td>
</tr>
<tr>
<td>Summary of the Study</td>
<td>119</td>
</tr>
<tr>
<td>Significance of the Study</td>
<td>120</td>
</tr>
<tr>
<td>Limitations</td>
<td>121</td>
</tr>
<tr>
<td>Research Question 1</td>
<td>123</td>
</tr>
<tr>
<td>Research Question 2</td>
<td>125</td>
</tr>
<tr>
<td>Research Question 3</td>
<td>127</td>
</tr>
<tr>
<td>Research Question 4</td>
<td>128</td>
</tr>
<tr>
<td>Recommendations for Practice</td>
<td>131</td>
</tr>
<tr>
<td>Recommendations for Future Studies</td>
<td>133</td>
</tr>
<tr>
<td>Conclusion</td>
<td>134</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>136</td>
</tr>
</tbody>
</table>

**APPENDIXES**

<table>
<thead>
<tr>
<th>A</th>
<th>Acronyms</th>
<th>145</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>Coding sheet</td>
<td>147</td>
</tr>
<tr>
<td>C</td>
<td>Studies included in the meta-analysis</td>
<td>154</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Percentage of Students Ages 6 Through 21 with Disabilities Served in Different Educational Environments during the 1997-98 School Year</td>
<td>10</td>
</tr>
<tr>
<td>2.</td>
<td>Various FBA Methods and the Corresponding Techniques with a Description of Each Technique</td>
<td>34</td>
</tr>
<tr>
<td>3.</td>
<td>Relationship Between Assessment Variables and Outcome Effectiveness</td>
<td>50</td>
</tr>
<tr>
<td>4.</td>
<td>Characteristics of the Studies Included in the Meta-analysis</td>
<td>102</td>
</tr>
<tr>
<td>5.</td>
<td>Summary of Cohen’s $d$ Index</td>
<td>106</td>
</tr>
<tr>
<td>6.</td>
<td>Average Effect Size for the Various FBA-based Methods and Interventions</td>
<td>108</td>
</tr>
<tr>
<td>7.</td>
<td>Average Effect Size for the Various FBA-based Methods and Interventions across Disability Types</td>
<td>109</td>
</tr>
<tr>
<td>8.</td>
<td>Analysis of Variance between Disability Type and Method of Assessment</td>
<td>111</td>
</tr>
<tr>
<td>9.</td>
<td>Average Effect Size for the Various FBA-based Methods Analyzed for Order</td>
<td>113</td>
</tr>
<tr>
<td>10.</td>
<td>Results of Effect Size Comparison by Model and Type of Functional Assessment</td>
<td>115</td>
</tr>
</tbody>
</table>
**LIST OF FIGURES**

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Percentage of students ages 6 through 21 in different education environments from 1988-89 through 1997-98</td>
<td>8</td>
</tr>
<tr>
<td>2. Four phases of functional behavioral assessment</td>
<td>12</td>
</tr>
<tr>
<td>3. Overview of the two-level FBA process</td>
<td>14</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

Statement of the Problem

Pervasive and maladaptive behaviors within public-education settings result in lower levels of academic engagement, grades, and performance on standardized achievement assessments (Shinn, Ramsey, Walker, O’Neill, & Steiber 1987; Wentzel, 1993). In practice, these behaviors are defined as self-injurious (e.g., self-mutilation, cutting, or head-banging), assaultive (e.g., severe verbal or physical aggression, possession of a firearm, or brandishing a knife on school grounds), or property damaging, which could lead to suspension or expulsion or could require frequent and systematic use of behavioral interventions (Code of Regulations, Title 5, Section 3001, Individuals with Disabilities Education Act; IDEA, 1997; Individuals with Disabilities Education Improvement Act; IDEIA, 2004; see Appendix A for a list of relevant acronyms). The terms pervasive and maladaptive refer to behaviors that are generally of an intensity, duration, and frequency that there is potential for (a) threat of harm to the individual, peers, or staff, (b) limited acquisition of novel skills or attainment of concepts by the individual or peers, (c) change of placement to a more restrictive environment, or (d) limited response to systematic interventions. To address these pervasive and maladaptive behaviors, the legislature (i.e., the House of Representatives, Congress, and governing state officials) developed policies and procedures based upon the principles of applied behavioral analysis (ABA), which are derived from the seminal works of Skinner (1938, 1953) and Bijou, Peterson, and
Ault (1968) in the form of positive behavioral intervention strategies and functional behavioral assessment (FBA).

The IDEA, reauthorized in 1997, and the 2004 IDEIA mandate that the relationship between academic achievement and disruptive behavior must be analyzed and addressed through the individualized education program (IEP) process. The 1997 amendment also included provisions stipulating that students with disabilities receive a free and appropriate public education (FAPE) in the least restrictive environment (LRE) to the maximum extent appropriate with other students who do not have disabilities. To ensure that students with disabilities receive academic and nonacademic educational benefit (e.g., socialization, extracurricular, and enrichment activities), the necessary supplementary aids and related services must be provided to address their unique needs.

As such, Section 614 (d)(3)(B)(i) of P.L. 105-17 or more commonly the IDEA 1997 states that when “behavior impedes his or her learning…the IEP team must consider, when appropriate, strategies, including positive behavioral intervention strategies” (p. 71). This section is proceeded by Section 615 (k)(1)(B)(i) that required that the local educational agency conduct an FBA to analyze the relationship between student behavior and the learning environment and implement a behavior intervention plan (BIP) before a change of placement due to numbers of suspensions or expulsions as a result of misconduct. An FBA is a systematic evaluation process for analyzing factors that may influence and maintain problematic behaviors through indirect, direct, or experimental methods of assessment (Sugai, Lewis-Palmer, & Hagan-Burke, 1999). The intended outcome of any FBA is to develop an efficient, effective, and relevant BIP that addresses undesirable target behaviors and fosters the
development of appropriate replacement behaviors (Sugai, Horner, et al., 1999). The assumption is that students with aggressive, assaultive, and even self-injurious behaviors can access the core curriculum within the LRE given a quality behavioral assessment and intervention.

The 1997 reauthorization of the IDEA, however, created several dilemmas for practitioners and researchers alike through the mandated use of FBA and positive-behavioral-intervention strategies in public-school settings. The following are some of the challenges noted in the literature and by practitioners in the application of the FBA process within public-school settings:

1. The FBA process and positive-behavioral-intervention strategies are empirically validated as effectively reducing the occurrence of problem behaviors and increasing the development of appropriate alternative behaviors. There is, however, limited support for the effective and efficient use of the FBA process within the public-school setting and even less evidence for students with the most prevalent types of disabilities (Carr et al., 1999; Ervin et al., 2001; Nelson, Roberts, Mathur, & Rutherford, 1999).

2. There is little agreement within the research community as to the most effective methods and procedures for conducting an FBA (Carr et al., 1999; Crone & Horner, 1999; Hanley, Iwata, & Mc Cord, 2003; Sasso, Conroy, Stichter, & Fox, 2001).

3. The terminology used to define the FBA procedures varies throughout the literature (i.e., functional analysis assessment, functional behavioral assessment, functional assessment, functional analysis, functional assessment of behavior, etc.; Cone, 1997).
4. The statutes do not provide procedural or practical guidelines for conducting the assessment for practitioners (Quinn, 2000; Sugai, Horner, et al., 1999).

5. ABA-based practices are a departure from the traditional school model that relies on contingent types of reinforcement, aversives, and punishment to address problem behaviors (Borgmeier, 2003, 2005; Mace, 1994).

6. Schools may not be equipped to address the needs of students with serious behavioral challenges, and there is a discrepancy between state and federal mandates and the resources, training, and ability to implement these assessment and intervention requirements effectively (Crone & Horner, 1999; Gresham, Watson, & Skinner, 2001; Horner, Diemer, & Brazeau, 1992).

7. Given the limited resources available to schools in terms of personnel, finances, and technical support, practitioners need clear guidelines to conduct an effective and efficient FBA to assist in the development and implementation of the behavioral intervention in order to bridge the research-to-practice gap (Crone & Horner, 1999; Horner & Sugai, 1999; Yell & Katsiyannis, 2000).

In practice, the term FBA refers to a variety of techniques and strategies to diagnose possible causes and identify likely interventions intended to address problem behavior (Quinn, 2000). These techniques include indirect (i.e., interviews, record reviews, rating scales, etc.) and direct descriptive methods (i.e., antecedent-behavior-consequence charting, time interval, intensity scales, etc.) to define operationally the behavior. Many practitioners use the term functional assessment or functional assessment of behavior to refer to any activity, which may include indirect and direct methods, involved in describing and formulating hypotheses about
potentially controlling variables (Cone, 1997). Other practitioners also use functional assessment to describe testing the hypothesis and evaluating the effectiveness of the intervention through the analysis of direct descriptive observation (Ervin et al., 2001). Compared with the research community, very few practitioners assist with or conduct experimental functional analysis assessments involving the testing or verification of hypotheses through the systematic manipulation of environmental features (Carr et al., 1999; Cone, 1997).

Many researchers viewed the reauthorization of the IDEA as an opportunity to move toward the application of applied behavioral analysis in classrooms, which would mean the development of assessment procedures, methods of training staff, and systems to monitor the effect of behavioral interventions. Ten years later, and with another reauthorization of the IDEA in 2004, the research community has maintained efforts to develop and analyze forms of sustainable methods for training staff and implementing PBS (Bergstrom, 2003; Carr et al., 2002; Horner, 2004). The need, however, persists in establishing guidelines for conducting an FBA in the school setting, utilizing local staff in order to realize these federal mandates (Borgmeier, 2005; Sugai, Horner, et al., 1999).

Therefore, this study analyzed recent single-case studies related to pervasive and maladaptive behaviors for students with high-incidence disabilities within the public-school setting. Given the wide-range of methods and interventions in the literature, a meta-analysis was conducted to evaluate the statistical and practical significance of each study in terms of substantive behavioral change between the baseline and intervention phase. This behavioral change was calculated as an average effect size for comparing treatment outcomes across studies. The meta-analytic
process allows the researcher to assign and weight values to studies in the form of
effect sizes for comparison (Rosenthal, 1991). In addition, these effect sizes were
examined for any differences in the methods of assessment or other relevant factors
such as disability type or the individual conducting the assessment or implementing
the intervention. A byproduct of this analysis may provide further guidance for a
more time and resource efficient FBA process and implementing effective behavioral
interventions within the public-school setting.

Background and Need

The development and implementation of FBA guidelines is influenced by a
number of factors that range from the national scale to the level of the individual
student. The first section of the background and need addresses the fundamental
changes in the federal and state laws that govern the educational rights for students
with disabilities and mandate the use of the FBA process. These changes in the law
have resulted in an increase in the number of students identified as needing special
education, a broader range of protections, and an enhanced level of inclusion into the
mainstream general-education environment.

Although FBA and positive behavioral interventions are recognized within the
research community, the practical application of ABA principles represents a
paradigm shift within the public schools. This applied approach to addressing
problem behavior represents a departure from a more punitive system to a more
positive-behavioral approach that emphasizes positive reinforcement of desired
behaviors and environmental manipulation as means in order to increase the
likelihood of their reoccurrence (Carr et al., 1999; Koegel, Koegel, & Dunlap, 1996).
The evolution of the FBA from the seminal works of B.F. Skinner in 1938 to more
current research is reviewed in the second section of the background and need. This longitudinal perspective lays the foundation for the use of current methods of FBA and interventions and identifies some limitations of the various FBA techniques.

The context of educating students with special needs. A number of prevailing factors influence the education of students with special needs on a national scale, particularly in the area of behavior. Changes in state and federal laws and levels of funding directly affect the guidelines related to the identification, discipline, services, and placement of students with special needs. In 1999, the year that the federal regulations for the IDEA 1997 were released, the Office of Special Education Programs (OSEP) established a goal that 50% of all school-aged children with exceptional needs were to receive 80% of their instruction in a regular-education setting. The OSEP goals paralleled the IDEA (1997) mandate. Students with disabilities must be educated to the maximum extent appropriate with their nondisabled peers. To protect the rights of students with special needs in inclusive settings whose behaviors are pervasive and maladaptive, the IDEA instituted, but did not define, FBA and PBS.

Meeting legal mandates often requires adding personnel, materials, and related services to a student’s program can be costly in terms of the requisite training and support to implement these behavioral practices. Since the 1997 reauthorization the IDEA, the federal government has allocated additional monies to train general- and special-education staff in effective inclusion practices and collaboration and in accommodating learners with diverse social, emotional, behavioral, linguistic, and academic needs. The actual federal P.L. 94-142 funding for the IDEA mandates, however, has increased since the 1997 reauthorization by only 0.6% to a level of
18.6% in 2005, making the IDEA an underfunded federal mandate (OSEP, 2006). Insufficiency of funds, however, is not grounds for denying a student with special needs a free and appropriate public education.

The concept of LRE as a pillar of the IDEA 1997 appears to have codified the prevailing trend toward increased levels of inclusion of students with special needs within the general-education setting (see Figure 1 and Table 1). Based upon a decade-long study by OSEP (2000), in 1998, 46% of school-age students with disabilities were served outside the regular classroom for less than 21% of the school day as compared with a level of nearly 31% outside the school day in 1989. Recent data from OSEP (2006) indicate that nearly 50% of students with special needs were served outside the regular classroom for less than 21% of the school day.

**Figure 1.** Percentage of students ages 6 through 21 in different education environments from 1988-89 through 1997-98. Figure adapted from the Office of Special Education Programs. (2000). *22nd Annual Report to Congress.* Data were derived from the Office of Special Education Programs, Data Analysis System (DANS).
As more students with special needs are included for longer periods of the day in the general-education setting, there is an increased need to develop effective methods for addressing behaviors in more natural settings such as the general-education classroom. More traditional methods involve removing the student from the learning environment or placing the student in a more restrictive setting outside of the regular class, which is in conflict with the notion of LRE. A student who is suspended or expelled from school in most cases cannot receive a FAPE. In an effort to ensure that special-education students are provided with FAPE, the policies related to FBA and positive behavioral supports were developed. This solution was chosen even with the limited research related to FBA procedures and interventions with high-incidence disabilities, such as learning disabilities, even though these students constitute the overwhelming majority of students served (Nelson et al., 1999).

Nearly 82% of students served with an IEP are eligible for services under at least one of the following categories: specific learning disability (SLD), speech and language impairment (SLI), emotional disturbance (ED), or other health impairment (OHI; OSEP, 2006). Table 1 contains data illustrating that students with high-incidence disabilities (SLD, SLI, or OHI) spend less than 60% of the school day outside of the mainstream setting for special-education services. Students with these types of disabilities spend the majority of their day in the general-education setting. Thus, there is a need to establish assessment practices and procedures in more naturalistic settings because the general-education setting has inherently higher levels of academic and behavioral expectations and greater student-to-staff ratio. These environmental factors in the general-education classroom often negatively influence student behavior if individualized strategies are not implemented; however, the
literature indicates that there are lower levels of staff involvement in the development and implementation of PBS and FBA (Carr et al., 1999; Reid & Nelson, 2002).

Table 1

Percentage of Students Ages 6 Through 21 with Disabilities Served in Different Educational Environments During the 2000-2001 School Year

<table>
<thead>
<tr>
<th>Primary disability</th>
<th>Served Outside the Regular Classroom</th>
<th>Separate environments a</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0-21% of the Day</td>
<td>21-60% of the Day</td>
</tr>
<tr>
<td>Specific learning disabilities</td>
<td>44.3</td>
<td>40.3</td>
</tr>
<tr>
<td>Speech or language impairments</td>
<td>85.6</td>
<td>8.4</td>
</tr>
<tr>
<td>Mental retardation</td>
<td>13.2</td>
<td>29.1</td>
</tr>
<tr>
<td>Emotional disturbance</td>
<td>26.8</td>
<td>23.4</td>
</tr>
<tr>
<td>Multiple disabilities</td>
<td>12.1</td>
<td>16.0</td>
</tr>
<tr>
<td>Hearing impairments</td>
<td>42.3</td>
<td>20.0</td>
</tr>
<tr>
<td>Orthopedic impairments</td>
<td>46.4</td>
<td>23.4</td>
</tr>
<tr>
<td>Other health impairments</td>
<td>45.1</td>
<td>33.9</td>
</tr>
<tr>
<td>Visual impairments</td>
<td>50.5</td>
<td>20.1</td>
</tr>
<tr>
<td>Autism</td>
<td>24.3</td>
<td>15.3</td>
</tr>
<tr>
<td>Deaf-blindness</td>
<td>18.1</td>
<td>9.9</td>
</tr>
<tr>
<td>Traumatic brain injury</td>
<td>32.3</td>
<td>27.9</td>
</tr>
<tr>
<td>Mental retardation</td>
<td>46.4</td>
<td>29.9</td>
</tr>
</tbody>
</table>

a Separate environments include public and private residential facilities, public and private separate facilities, and homebound/hospital environments.

Note. Table adapted from the Office of Special Education Programs. (2006). 25th Annual Report to Congress. Data were derived from the Office of Special Education Programs, Data Analysis System (DANS).

Even though the large number of students with high-incidence disabilities and their high level of involvement in the mainstream setting, these students are underrepresented in the literature with regard to the development and implementation of FBA and PBS (Dahlstrom, 2003; Nelson et al., 1999). Therefore, this study intends to build on the existing body of research that focuses primarily on FBA and positive behavioral interventions with students and adults with low-incidence disabilities such
as mental retardation or autism by evaluating the various techniques of behavioral assessment and interventions with high-incidence disabilities in public-school settings. This second section of the background and need examines the evolving literature base, policy, and practice related to FBA.

*Functional behavioral assessment.* Functional behavioral assessment is the systematic study of causal or related factors contributing to the occurrence of behavior as a means to develop effective interventions (Gresham et al., 1999). The analysis involves examining the relationship between setting events, antecedents, and consequences that interact with disruptive behavior. The FBA process, based upon the work of Skinner (1938, 1953) and Bijou et al. (1968), provides educators and researchers with the means to operationalize behaviors in terms of stimuli that trigger a behavior and appropriate responses to reinforce desired behavior. Contemporary schools more commonly utilize arbitrary contingencies of reinforcement and punishment such as detention, suspension, and expulsion instead of individualized treatments that are based upon an understanding of the function of the behavior, (Mace, 1994).

Functional behavioral assessment is an evolving technology grounded in the principles of applied behavioral analysis (Carr et al., 2002). There is widely-accepted agreement that a FBA is conducted to (a) determine variables that relate to the occurrence and nonoccurrence of behavior, (b) identify environmental factors that may be manipulated to influence behavior, and (c) define operationally the functional relationship between the behavior or class of behaviors for an individual (Hanley et al., 2003; Iwata et al., 1994; Sugai, Horner, et al., 1999). Collectively, these data from the FBA are used to establish a conditional or predictive probability of behavior in
order to develop interventions and supports that result in a decrease of target behaviors and an increase of appropriate alternative behaviors (Dunlap, Dunlap-Kern, Clarke, & Robbins, 1991; O’Neill et al., 1997; Reid & Nelson, 2002; Yell & Katsiyannis, 2000). Even though there are foundational agreements within the literature, there are differing perspectives on effective and efficient procedures and models for conducting an FBA and implementing the intervention plan.

To analyze FBA in school settings, Ervin et al. (2001) examined 100 studies conducted with disabled and nondisabled students across eligibility categories and ages. This comprehensive critique differentiated four procedural phases of behavioral assessment and intervention (see Figure 2). To be included in the synthesis, the study had to include at least one of the phases or any combination of the four. The first descriptive phase uses either indirect (e.g., surveys, interviews) or direct methods (e.g., descriptive observations) to assist in the identification of related variables and behaviors.

![Figure 2. Four phases of functional behavioral assessment.](image-url)
The second is an interpretive phase in which a hypothesis is developed based upon the perceived relationship of the variables to the behavior. The third phase involves the formal testing of the hypothesis. The final phase involves the implementation of the intervention. This review of the literature by Ervin et al. (2001) found that 142 of the 148 intervention cases were successful in decreasing negative behaviors within the school setting.

The study by Ervin et al. (2001) presents two points that are very relevant to the current study. The first is that an FBA that consists of various methods or phases (i.e., descriptive, hypothesis, and implementation) can yield behavioral change. The second important aspect of this recent study is that the researchers present a model for conducting an FBA that involves an evaluation of the behavior and the intervention. The difficulty for future synthesis of the literature in evaluating this model is that the assessment procedures within the various phases are often nonspecific.

Other researchers have promoted a two-level FBA approach that can be modified according to the intensity and nature of the behavior (see Figure 3; Sugai, Lewis-Palmer, et al., 1999). With this approach, the level of intensity of the behavior determines the level of resources required to conduct the assessment and likely implement the intervention. In practice, the evaluator has the flexibility to conduct the minimal amount of assessment for the optimal effect. If the intervention is not successful, then the practitioner conducts a more time and energy intensive functional assessment to evaluate the intervention. The functional hypothesis is tested and measured through the success of the intervention plan. An ineffective plan warrants a further, more intensive assessment that examines the conceptualization of the behavior as well as other factors that can influence the effectiveness of the
implementation such as the teacher’s willingness and ability to implement the plan, reinforcement procedures, and so forth. Progress monitoring increases the fidelity and efficacy of the assessment and intervention. The flexibility inherent in the multiple-phases model (Ervin et al., 2001) and the two-level FBA model (Sugai, Lewis-Palmer, et al., 1999) provides researchers and practitioners with a framework upon which indirect, direct descriptive, and functional assessment methods can be examined.

Figure 3. Overview of the two-level FBA process. Figure adapted from “Overview of functional behavioral assessment process.” G. Sugai, T. Lewis-Palmer, and S. Hagan-Burke, 1999, Exceptionality, 8, p.156.

The initial work begun by Iwata et al. (1982) to structure a comprehensive replicable model of functional assessment has led to over 700 published studies based
upon this assessment process (Hanley et al., 2003). Although there are an abundance of case studies and procedural analyses, there is little agreement as to what constitutes a comprehensive functional behavioral assessment. Some researchers dispute whether nonexperimental methods of behavioral assessment, such as descriptive analysis or interviews, are adequate when compared with experimental approaches such as a functional analysis that test the hypothesis through the systematic manipulation of environmental events (Doggett, Edwards, Moore, Tingstrom, & Wilczynski, 2001; Gresham et al., 2001; Horner & Carr, 1997). Although the settings, behavior, samples, and other factors may vary, the foundation of applied behavioral methodology remains the same: identify environmental and interpersonal factors and the communicative intent related to a given behavior in order to teach alternate behaviors more effectively.

*The evolution of FBA.* Skinner (1938) explained functional analysis as the empirical demonstration of cause-and-effect relationships between the environment and behavior. This relationship between behavior and the environment is the basis for the psychological theory of behaviorism and operant learning. Skinner analyzed these contingent relationships to ascertain a means for manipulating the environment to change an individual’s behavior.

In a later essay, Skinner (1953) included the notion of antecedents and consequences as relational factors that could influence behavior. This seminal work established the contingencies of positive and aversive stimuli or the provision or removal of something perceived as positive or aversive. Such environmental manipulations allow the researcher or practitioner to predict more easily the occurrence or nonoccurrence of the behavior. Although there often was a clear causal
relationship, consequences such as response cost, time out, tangible items, or preferential and non-preferential activities were often unsuccessful because of the limited understanding of other factors influencing behavior (e.g., multiple-functions of behavior, etc.). Antecedents were defined as events occurring before a behavior, which are setting events, establishing operations (EO), and discriminative stimuli (SD). An EO temporarily alters the effect of a reinforcer for a given behavior, whereas SD suggests that a particular behavior will occur based upon past responses or schedules of reinforcement. Setting events are the farthest removed in terms of time and location from the behavioral incident yet are still related functionally (e.g., a student having a disagreement with a sibling before school or missing his or her medication).

The majority of early-functional analyses, which were aligned with Skinnerian principles, involved the relationship between self-injurious behavior (SIB) and differential reinforcement, depending upon the individual (Bachman, 1972; Carr, 1977; Lovaas, Freitag, Gold, & Kassorla, 1965). Typically, the research conducted during this period occurred in clinical settings within institutions, and the participants were primarily individuals with mental retardation or with mental retardation and a concurrent diagnosis of autism. This early research led to the conceptual establishment of function or why a behavior was occurring (Bijou et al., 1968).

Lovaas et al. (1965) introduced the effects of social-positive reinforcement or attention on a population dually diagnosed with mental retardation and autism within clinical settings. In an effort to examine whether social-positive reinforcement generalized to other settings and disabilities, other studies conducted within the classroom setting provided promising implications in working with students with
aggressive (Pinkston, Reese, Le Blanc, & Baer, 1973) and disruptive behaviors (Thomas et al., 1968). Shortly after the establishment of attention as an operant reinforcement, researchers recognized that behavior also was maintained through negative reinforcement or escape from a stimulus, such as a task or instruction that was perceived as being too difficult (Sailor et al., 1968; Weeks & Gaylord-Ross, 1981).

Similarly, Carr, Newsom, and Binkoff (1980) expanded the escape function to include the removal or presentation of certain stimuli as a correlate to SIB and aggressive behavior. If attention and escape functions of problematic behavior could be identified, then these same behaviors could be reduced through systematic ABA interventions (Carr et al., 1980).

In a single-case study, Iwata et al. (1982) used previously established concepts to develop a comprehensive functional analysis for SIB that entailed direct observation and repeated measurements applied across four conditions (three treatments and one control) in this single-case study. The three treatments involved establishing operations, discriminating stimuli, and identifying reinforcements for each contingency. The control was without the aforementioned conditions. Although the levels and functions of the SIB varied, the results of this research suggested a statistically and practically significant reduction in response to the treatments (Iwata et al., 1982).

The establishment of the functional analysis method, as a more positive form of behavioral support that emphasized evaluation and understanding of behavior to teach alternatives, resulted in the reduction of more aversive techniques and punishment (Pelios, Morren, Tesch, & Axelrod, 1999). Observations of the U.S.
penal system, as well as governmental sources, have suggested that aversive practices are more effective in reducing pervasive and maladaptive behaviors as compared to applied-behavioral-analytic approaches (U.S. National Institutes of Health, 1989). The dramatic increase in the number of prisons built on a national level suggests that aversive approaches are ineffective. Furthermore, the high recidivism rates in the penal system suggest continued research and education in the area of applied behavioral analysis and positive-behavioral supports as a preventative solution that teaches alternatives to antisocial and criminal behaviors (Carr et al., 1999; Hanley et al., 2003).

As the FBA process has evolved, researchers and practitioners alike sought a means to develop a process that was time-and-cost-effective utilizing the best practices from nearly 70 years of research. This refinement of the assessment has resulted in three methods: indirect, direct, and functional assessment.

Emergence of the multiple methods of FBA. Functional behavioral assessment has provided researchers and educators alike with a systematic procedure to evaluate which interventions may be more effective given an identified function of the behavior (Carr et al., 1999; Gresham et al., 2001; Hanley et al., 2003). This assessment process uses multiple methods for analyzing the relationship between behavior and the environment in operationalizing the hypothesis of the target behavior. The functional assessment evaluates the accuracy of the hypothesis and in some models the effectiveness of the intervention.

The first is the indirect method, which can often be used to gain initial insight into problematic behaviors through a review of archival records, rating scales, checklists, or interviews (Fischer, 2003; Gresham et al., 2001; Horner & Carr, 1997).
The second method, involves direct descriptive observations that can be used to help formulate the hypothesis for the occurrence of the behavior and quantify behavioral events through various methods of data collection (e.g., frequency, temporality, intensity, or permanent byproduct techniques; O’Neill et al., 1997).

For the purpose of this study, the last method will be referred to as a functional assessment (FA) and is defined more broadly as any method of assessment that evaluates the accuracy of the hypothesis and, in many cases the effectiveness of the intervention (Carr et al., 1999; Gresham et al., 1999). This broad interpretations is supported through a comprehensive review of the literature that revealed a wide range of terms used to describe this method within the FBA process, such as functional assessment, functional analysis, functional analysis assessment, brief functional assessment, behavioral assessment, and functional communication (Carr & Durand, 1994; Carr et al., 1999; Henley et al., 2003; Horner, 1994; Iwata et al., 1994; Northrup et al., 1991; Sasso et al., 2001). These terms are often used interchangeably within the literature and the legislation, which can create further confusion for practitioners (Cone, 1997; Ervin et al., 2001).

Methodologically, FBA continues to evolve as researchers in the field of applied behavior analysis refine the standards for conducting these assessments. This section examines each of the methods and their supporting techniques in terms of practical application and the possible limitations.

In light of the limited level of training available, qualified staff needed to conduct the assessments, increased demand for the assessments, and the considerable amount of time required to conduct an FBA, the research and educational community
are examining more efficient, effective, and acceptable methodologies for use in the public-school setting.

Dahlstrom (2003) analyzed the psychometric properties of FBA by dividing the assessment into seven possible permutations: Indirect only, Direct descriptive observation only, Functional Assessment (FA) only, Indirect + Direct descriptive, Indirect + FA, Direct descriptive + FA, and Indirect + Direct descriptive + FA. An FBA case study was developed by the University of Oregon School Psychologist Project and was presented to a sample of 625 practicing school psychologists affiliated with the National Association of School Psychologists (NASP). Participants were given case studies based on one of the seven methods of conducting an FBA. Participants were limited to a single method by the corresponding data they received, that is, the Indirect only group was given a summary of antecedent, behavior, and consequence data in an archive review and the Functional Analysis Inventory (FAI; O’Neill et al., 1997). The participants were to identify antecedent events and maintaining consequences based upon the information derived from one of the various FBA methods. In this study, the dependent variables were the identification of the behavioral conditions (i.e., setting events, antecedents, consequences, etc.) and formulation of an accurate hypothesis. The independent variables were the method of FBA and the level of training for each respondent. A two-way analysis of variance (ANOVA) was used to compare the accuracy of the individuals in identifying antecedents and consequences across the various FBA procedures. The results of the study suggested that there was no statistically or practically significant difference between the various methods of FBA.
Although the results were not significant statistically, the results provide a model for analyzing the fundamental components of an FBA. These methods and the various permutations of the FBA methods are relevant in that the current study is intended to examine the relative effects of these methods within the existing body of behaviorally-related literature.

Indirect assessments or indirect procedures, such as archival-record reviews, questionnaires, behavioral rating scales, environmental scales, and interviews can be effective instruments in identifying possible antecedents, consequences, and in determining individualized reinforcements or schedules (Gresham, Watson, & Skinner, 2001; Mace, 1994). Indirect assessment is beneficial when analyzing behaviors that are difficult to observe including low-frequency, high-intensity behaviors such as possession of a controlled substance or weapons possession. Furthermore, these descriptive measures can be useful in identifying setting events or more remote antecedents (Carr, 1994). Similarly, this method can aid in evaluating students who are unable to attend school as the result of a suspension or expulsion or have been placed on an interim alternative educational placement. These relevant yet often overlooked factors that impact behavior may include psychological or physiological states related to drug exposure, illness, trauma, changes in the family unit, social changes, or sleeping and eating patterns.

The functional assessment interview (FAI) by Horner and Carr (1997) is one of the most frequently and widely used forms of indirect data collection. The assessor begins the process by interviewing individuals most directly associated with the student such as parents and teachers and, when appropriate, the individual exhibiting the target behaviors (O’Neill et al., 1997). The purpose of the FAI is to define
operationally the behavior, identify antecedent conditions, propose a hypothesis related to the suspected function of the behavior, and suggest possible alternative or replacement behaviors that meet the function of the behavior (Gresham et al., 2001; Horner & Carr, 1997; O’Neill et al., 1997). Utilizing this instrument allows the evaluator to collect information from multiple sources and contexts. This information guides the recommendations for the next course of action whether that is further data collection or the development of an intervention plan.

The limitations of indirect methods are the subjective nature of this form of FBA, which makes this type of assessment the least reliable of the three methods. The results of this method often reflect the perception of a few individuals, which may or may not be an accurate or objective depiction of the behavior and the maintaining variables (Horner & Carr, 1997). Moreover, indirect measures warrant further empirical research as these assessments provide valuable insight into the more idiosyncratic elements of behavior (Doggett et al., 2001; O’Neill et al., 1997). Gresham et al. (2001) and Ervin et al. (2001) have suggested that indirect methods serve as the initial phase or step that guide the assessor toward a more comprehensive FBA.

The second method for conducting an FBA is direct descriptive observation, which entails going to the setting or settings where the behavior is most likely to occur (Ervin et al., 2001). Whereas the first indirect method is removed from the behavioral event, descriptive observations provide in vivo data as to the frequency, temporality, intensity, and physical byproducts of the problem behavior (Fischer, 2003; Horner & Carr, 1997). Additionally, the use of more standardized formats of
documenting antecedents, behaviors, and consequences has become a mainstay of FBAs in recent years (Hanley et al., 2003).

Of 536 graphed individual data sets of functional analysis outcomes reviewed in the Hanley et al. (2003) study, 514 or 96% reflected effective treatments across a variety of singular behavioral topographies, such as SIB, aggression, disruption, property destruction, pica (i.e., the abnormal oral consumption of objects not suitable for eating), and vocalizations. Of the 514 studies reviewed, 87% utilized antecedent-behavior-consequence (ABC) models to assess behaviors (Hanley et al., 2003; Iwata et al., 1994). Alternately, Carr and Durand (1985) proposed the antecedent-behavior (AB) model in a study that examined a reduction of problem behaviors through instruction in communicative strategies. This study suggested a positive reduction of the target behaviors; in other words, the AB method may be a more time efficient means of assessing antecedent-based protests related to communicative deficits. Hanley et al. (2003) reported that 20.2% of the studies employed the AB technique, which is unusual given the lack of stimulus control in this method of assessment.

Other techniques used in conducting direct observations include the following:

1. Event-based recording documents the exact frequency of the behavior. This form of documentation is used best with behaviors that have a clear beginning and end and that occur at a mild to moderate rate (Gresham et al., 2001; Horner & Carr, 1997; O’Neill et al., 1997).

2. For more involved behaviors that are more ongoing in nature, the interval or time-based method can assist the assessor in intermittently establishing
whether a behavior occurred during the observation period (Gresham et al., 2001; O’Neill et al., 1997).

3. Observations of temporal characteristics of behavior are helpful in assessing the duration of a behavior or the latency for a response to a stimulus. Duration involves recording the length of behavioral event. Latency recording refers to the amount of time before an individual responds to some kind of environmental input (Gresham et al., 2001; O’Neill et al., 1997).

4. Permanent byproduct and intensity measures are used less frequently and can be somewhat problematic (Gresham et al., 2001).
   
   a. Documenting permanent byproducts, such as broken items, requires little effort. Identifying the responsible individual is sometimes more difficult (Dahlstrom, 2003).
   
   b. Intensity involves collecting data with regard to the severity of behavior. Often the measure of intensity is a synthesis of frequency, duration, and permanent byproduct data. Even with reasonable documentation, however, deciding what constitutes a moderate or severe behavior is somewhat subjective. Over time, the observer and individuals familiar with the subject may be able to create individualized descriptors or rubrics to define the intensity of a behavior or class of behaviors.

Although direct descriptive observations can provide observers with relevant data in proposing a hypothesis for the function of a behavior or class of behaviors, there are limitations to consider. As mentioned above, this method can be difficult to standardize across observers, environments, and even behaviors, which poses a threat to the validity and interrater reliability of the measure. Moreover, the data are
correlational, which can suggest functional rather than causal relationships between
the behavior and the environment (Mace, 1994; Mace, Yankanich, & West, 1989). Of
the three methods, an experimentally designed functional assessment is the only
method that can demonstrate a causal relationship to a statistically and practically
significant degree (Hanley et al., 2003, Iwata et al., 1994).

Within the category of FA is functional analysis, which is the experimentally
controlled manipulation of environmental factors to assess their affect on the target
behavior (Hanley et al., 2003). These responses to the environmental controls are
used to develop a hypothesis, which can then be tested through further exposure to
the relevant conditions (Gresham et al., 2001).

Through their extensive synthesis of the literature, Hanley et al. (2003)
proposed recommended practices in developing function-based analysis and
interventions that include (a) limiting the number of behaviors being analyzed, (b)
testing to investigate the relationship between consequences and the target behaviors,
(c) identifying establishing operations prior to and during the evaluation, (d)
accounting for discriminative stimuli, (e) limiting session length, (f) testing for a
functional relationship between behavior and tangible reinforcer when appropriate,
(g) accounting for relative reinforcement duration in analysis of results, (h) beginning
assessment with brief and basic trials then graduate to more complex and lengthy
tasks, and (i) utilizing supplementary assessments in preparation and in conjunction
with more formalized procedures.

Perhaps the greatest challenge to generality of a functional analysis is that
most are conducted in controlled or analog environments (e.g., clinics, hospitals, and
institutions; Gresham et al., 2001). These controlled environments may be different
from the individual’s natural environment, such as in a classroom where the behavior is more likely to occur. Critics of analog assessments conducted within a clinical setting believe that because the target behaviors are decontextualized, analysts may not be able to determine potential antecedents and consequences, which occur only in the natural setting (Gresham et al., 1999). Many behaviorists, however, have argued that clinical settings enable the analyst to influence variables more effectively and reduce the number of extraneous stimuli, which following this line of reasoning, allow for a higher level of confidence in the findings (Hanley et al., 2003).

In practice, there are many threats to the validity and reliability of the FBA process. Researchers have cited the complicated and time-intensive nature of conducting the assessment as a primary threat to the acceptability and feasibility of FBA within the public-school setting (Gresham et al., 1999, 2001; Horner & Carr, 1997). Moreover, educators are faced with balancing the initial investment of time to acquire the requisite skills and to conduct a FBA within the classroom environment with actually teaching the class, completing report cards, as well as maintaining positive relationships with students, families, and staff (Bergstrom, 2003; Crone & Horner, 1999). Studies in the area of teacher preparation and implementation of research-based behavioral practices cited limited training in behavior-management techniques and analysis as a factor that impedes appropriate functional behavioral assessment (Horner et al., 1992; Kaufman & Wong, 1991; Scott & Nelson, 1999; Watson & Robinson, 1996).

Because of limited experience, time, and resources, the educational system often relies on school psychologists to perform the functional behavior assessment (cf., Gresham et al., 2001). Based upon estimates from the Northern California
Diagnostic Center in conjunction with the Positive Environments Network of Trainers (2003), school psychologists who are well versed in the FBA process can complete a comprehensive functional behavioral assessment in 16 to 30 hours following standardized procedures that include archival review, indirect, direct, and brief experimental analysis. A trained consultant with a background in applied behavioral analysis can complete a similar FBA in 9.7 to 23 hours (Schill, Kratochwill, & Elliott, 1998). These limitations of the FBA process, however, have resulted in researchers developing alternative forms of the traditional FA.

In an effort to streamline the FA process and to make it more practitioner friendly, researchers have developed variations commonly referred to in the literature as brief functional assessment or analysis approach, depending upon the researcher (Gresham et al., 2001; Northrup et al., 1991). Most brief FAs consist of a short analog assessment of two or more conditions and then a replication of these conditions in a more natural environment (Northrup et al., 1991). Sigafoos and Saggers (1995) utilized short time intervals of 2 minutes for 20 trials within a more naturalistic setting. This study examined the effect of two one-minute contingencies where the reinforcement was presented intermittently across the subject’s school day. The result of the Sigafoos and Meikle (1996) study, a replication of the 1995 short-interval study, suggested differential responding in both instances. These two studies established the brief functional assessment with more complex environmental variables (e.g., schedules, other students, noise, and movement typical elements of a regular classroom) as a viable alternative to more involved functional analyses that follow a rigorous experimental design.
Although there is a relatively small investment of time to conduct a brief FA, it is not without limitations. This practice requires an individual who is versed in applied behavioral analysis and has the ability to observe unobtrusively in a complex environment, such as a classroom (Hanley et al., 2003; Sugai, Horner, et al., 1999).

Because of the time required to conduct multiple observations and the intrusive nature of observing within a natural context, it may be difficult to establish consistently valid and reliable observations across activities and environments (Northrup et al., 1991).

In summary, there are various methods of conducting a FBA with a wide range of strengths and limitations (Dahlstrom, 2003; Gresham et al., 2001; Hanley et al., 2003). Furthermore, there is support that these methods and techniques may provide educators with an effective way to examine the complexities of behavior and to devise interventions to address pervasive and maladaptive behavior in the public-school setting (Carr et al., 1999, 2002; Doggett et al., 2001). Although there is overwhelming support of the FBA process in the literature, there is a gap among legal mandates, the research community, and practitioners with regard to recommended best practice guidelines for assessment, development and implementation of intervention plans, and training models for staff. Current research in the field focuses primarily on the methods of FBA or specific interventions. When studies do analyze the relationship among an assessment and the resulting intervention, there generally is little comparison between similar studies. Therefore, the purpose of this study was to utilize existing studies to analyze methods of FBA, types of interventions, and the resulting effects to suggest more efficient and effective methods of conducting behavioral assessments within the public-school setting.
Purpose of the Study

The purpose of this study was to conduct a meta-analysis of single-case studies of students with high-incidence disabilities attending public schools exhibiting pervasive and maladaptive behaviors. First, leading journals in the fields of behavioral analysis, positive behavioral support, psychology, school psychology, and special education, as well as related dissertations and professional articles, were reviewed for potential sources for studies involving school-aged students with high-frequency disabilities. References of the obtained articles were reviewed for additional sources. Material found through this search was examined further in terms of the characteristics of the student, the topography or behavior in operational terms, the setting of the assessment, the function of the behavior, the method of analysis, the type of intervention, and the outcome of the intervention. Once this information was categorized, a meta-analysis was conducted to compute the effect size of the behavioral change between the baseline and intervention phases for comparison. The studies then were analyzed in terms of the FBA procedure used in developing the intervention based upon one of the following methods: Indirect only, Direct descriptive observation only, Functional Assessment (FA) only, Indirect + Direct descriptive, Indirect + FA, Direct descriptive + FA, and Indirect + Direct descriptive + FA. Additionally, the multiple-phase and two-level models that guide the assessment and intervention process were evaluated (i.e., Ervin et al., 2001; Hanley et al., 2003; Sugai, Lewis-Palmer et al., 1999). The comprehensive review of the literature and the findings of this meta-analysis may assist in establishing practical guidelines for practitioners and in bridging the research-to-practice gap in the
development and implementation of mandated FBAs and BIPs (Reid & Nelson, 2002).

Research Questions

This study addressed the following research questions:

1. What is the average effect size for each of the various FBA-based methods and interventions for students with high-incidence disabilities, such as learning disabilities, attention deficit disorder with hyperactivity, or emotional/behavioral disorders?

2. Which FBA methods yielded the greatest average effect sizes, and how did the various methods compare with each other across the various high-incidence disabilities?

3. If multiple FBA methods were used, was there a difference in the effect sizes based on the order in which the methods were used?

4. Did the assessment and intervention process follow a specific framework, such as a multiple-phase or two-level model, from beginning to end? If so, was there a difference in effect sizes for the interventions developed within the given framework?

Significance of the Study

Federal and state mandates have placed additional responsibility on school staff in developing interventions that ensure students receive a free and appropriate public education in the least restrictive environment while enhancing students’ academic achievement and social-emotional development. This study was intended to assess the overall effectiveness of FBA-based interventions in terms of the reduction of pervasive and maladaptive behavior and an increase of appropriate behaviors in the
public educational settings. The reduction of such behaviors may affect levels of active engagement in student learning, increased teacher satisfaction and perceptions of self-efficacy, as well as the development of prosocial behaviors in students. In order for these behavioral changes to occur within the public-school setting, appropriate and relevant models and methods for conducting a FBA must be developed.

Educational practitioners may find current studies difficult to translate into practice because of the varied settings, conditions, and student populations. To bridge this research-to-practice gap, this study explored the application of this technology in public-school settings with representative samples using methods that incorporate local educators in the assessment and implementation process. The development of a systematic process for conducting an FBA may increase the likelihood of the development of effective interventions. To analyze the relationship between the method or methods of assessment and the intervention, the effect sizes for the selected studies were compared by method of assessment. Additionally, it was beneficial to examine the order of multiple methods used in the FBA, for example, did the study employ an archival review to determine which direct descriptive analysis to pursue.

*Theoretical Rationale*

The most recent reauthorizations of the Individuals with Disabilities Act of 1997 and 2004 utilized an ample body of research from nearly a century of work in the field of behaviorism in the mandated application of FBA and BIP in the public-school system. Despite many misconceptions about behaviorism, this theory has evolved greatly since the seminal works of B.F. Skinner. Skinner’s (1953) operant
behavior theory provided the research community with a model to depict the relationship between behavior and the environment. These contingent relationships could then be further examined in terms of relationships that, when manipulated, influence learning.

This influential understanding of contingencies developed further into applied behavior analysis as researchers sought to translate behavioral theory into a more practical context (Baer, Wolf, & Risley, 1968). Early applied behavior analysts relied upon interventions that manipulated consequences. Bijou et al. (1968) proposed that there might be an underlying function to behavior that communicated a certain need. By identifying the function of behavior, more appropriate interventions could be selected. The most commonly identified functions of behavior are (a) get or obtain or (b) escape or avoid a wide-range of positive and aversive stimuli such as attention, internal events (e.g., automatic reinforcement, onset of a seizure, hyper- or hypo-arousal), replication of a chain of behaviors, and so on (Carr, Yarborough, & Langdon, 1997). An example of this functional relationship might be a student who is having difficulty with a mathematics lesson and whistles loudly to gain the attention of the instructor. Without an understanding of the function, a teacher might remove the student from the class, which might result in the student not getting the attention and assistance needed to complete the task. This understanding of the relationship between the function of behavior and consequences ultimately distinguished FBA from other methods of psycho-educational assessment.

As further methods of FBA emerged and were refined, several frameworks evolved to define the relationship between the assessment process and the interventions that resulted from the analysis of student behavior. Ervin et al. (2001)
identified four phases that characterized the assessment and implementation process (see Figure 2). The alternate method, the two-leveled approach of conducting a FBA and developing interventions, found in Figure 3, involves an assessment that becomes more complex and involved commensurate to the intensity of the behavior (Sugai, Lewis-Palmer, et al., 1999). Both models share common methods of assessments that include indirect, direct descriptive, and functional assessment (O’Neill et al., 1997).

Illustrated in Table 2 are a variety of techniques within each FBA method that can be employed in conducting an FBA. There are several assumptions that are consistent throughout the various methods. The first assumption is that nearly all sources of behavior are contained within the immediate observable environment, with perhaps setting events as the exception (Schill et al., 1998). In other words, assessed behaviors occur within and in response to distinct environmental situations and responses. Moreover, the conclusions of an assessment are neither generalizable to settings where the behavior has not been observed nor applied to other individuals (Gresham et al., 2001; Schill et al., 1998). Due to the nature of the data collected, whether through record review, direct observation, or a functional assessment, each assessment is unique to the subject. The final assumption is that behavioral assessments do not address internal factors that may be associated with the occurrence of the behavior such as depression or a personality disorder.
Table 2
Various FBA Methods and the Corresponding Techniques
With a Description of Each Technique

<table>
<thead>
<tr>
<th>FBA Method</th>
<th>Technique</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indirect</td>
<td>Archival review</td>
<td>Review of cumulative, previous test data, anecdotal teacher reports, behavioral logs, IEPs, medical and health records, and other relevant documentation.</td>
</tr>
<tr>
<td></td>
<td>Rating scales and Checklists</td>
<td>May examine environmental, behavioral, academic, preference, reinforcement, perceptions, related to antecedents, behavior, and consequences (ABC).</td>
</tr>
<tr>
<td></td>
<td>FAI (Horner &amp; Carr, 1997)</td>
<td>Assists in identifying the behavior operationally as far as when and where a behavior is more or less likely to occur, potential reinforcers, reactions and responses from others, possible functions, and hypotheses.</td>
</tr>
<tr>
<td></td>
<td>FACTS (Borgmeier, 2003;Marche et al., 2000)</td>
<td>Interview tool that assists in the development of hypotheses by examining function and ABCs.</td>
</tr>
<tr>
<td>Direct descriptive observation</td>
<td>AB (Carr &amp; Durand, 1985)</td>
<td>Antecedent-behavior model documents what happens before the behavior occurs.</td>
</tr>
<tr>
<td></td>
<td>ABC (Iwata et al., 1994)</td>
<td>Antecedent-Behavior-Consequence model documents what happens before, during, and after a behavioral episode.</td>
</tr>
<tr>
<td></td>
<td>Frequency</td>
<td>How often a behavior occurs or event-based recording during set intervals.</td>
</tr>
<tr>
<td></td>
<td>Temporality</td>
<td>When the behavior occurs or time-based methods that may include duration, latency, intensity.</td>
</tr>
<tr>
<td></td>
<td>Intensity</td>
<td>How severe is the behavior.</td>
</tr>
<tr>
<td></td>
<td>Permanent byproducts</td>
<td>Records physical products resulting from behavior, for example, property destruction a broken pencil or torn assignments.</td>
</tr>
<tr>
<td>Functional assessment</td>
<td>Experimental manipulation (Iwata et al., 1994)</td>
<td>Examines the relationship between the behavior and the environment in which the hypothesis is tested through exposure to multiple conditions in an analog environment.</td>
</tr>
<tr>
<td></td>
<td>Brief FA (Northrup et al., 1991)</td>
<td>Examines the relationship between the behavior and the environment in which the hypothesis is tested through exposure to two or more conditions briefly in an analog setting. The test is replicated with the conditions prompting least and most appropriate responses in a natural context.</td>
</tr>
</tbody>
</table>

Classical behaviorism and the tenets of ABA serve as the theoretical foundations for PBS and FBA as viable strategies in public schools. There is, however, a lack of consensus within the research community as to what constitutes a FBA in practical terms. Given the presence of state and federal mandates and the absence of clear guidelines for the application of ABA principles, practitioners are faced with a research-to-practice gap. To address this dilemma, this meta-analysis examined the models that give the assessment methods shape in a procedural sense.
and the effectiveness of the intervention, which is a measure of the quality of the assessment process, in an effort to provide guidance to practitioners.

*Definition of Terms*

Given the various uses and interpretations of behaviorally-based terminology, this section addresses terms as they apply to this study. Furthermore, this section uses definitions from the current literature that are most likely to be encountered by the practitioner.

**Antecedent** refers to the event that immediately precedes a behavioral event or class of behaviors (Bijou, Peterson, & Ault, 1968; Hanley et al., 2003). This event can be used in evaluating the conditional probability that a given behavior may occur. In terms of environmental manipulation, these conditions are the most easily modified to increase the likelihood that appropriate behavior or replacement behaviors may occur (Ervin et al., 2001). Antecedent conditions are particularly relevant to this study as a critical component in indirect, direct, functional assessment, and combined methods of FBA (Dahlstrom, 2003).

**Applied behavioral analysis (ABA)** is based upon radical behaviorism of B.F. Skinner and is the systematic approach to the assessment and the application of interventions that alter behavior (Baer et al., 1968). The FBA methods and interventions examined in this study adhere to the learning theory and the application of this theory, which is commonly referred to as ABA.

**Atypical intervention agent** is an individual whose primary purpose for interacting with the individual is to implement the intervention plan such as a clinical behaviorist, researcher, research assistant, and so on (Carr et al., 1999). As this study
did not specifically address the effects of typical versus atypical intervention agents, studies involving atypical agents were included in this meta-analysis.

*Atypical intervention setting* refers to a location where the intervention plan is implemented that is considered a less natural setting, such as a behavioral clinic, residential setting, hospital, and so on (Carr et al., 1999). Research with results obtained purely in atypical settings were excluded from this meta-analysis.

*Behaviorism* is the psychological model that suggests that behavior is governed by the relationship between the individual and his or her environment (Gresham et al., 2001). In this research, the assessment process seeks to establish hypotheses that illustrate the interaction between the individual and his or her environment, which results in interventions that may be tested and refined through further analyses. Contrasting theories suggest that behavior is motivated internally instead of as a result of environmental conditions that shape behavior.

*Behavior intervention plans (BIP)* are developed to address target behaviors through (a) the systematic instruction of alternative replacement behaviors that meet a similar function, (b) the manipulation of environmental conditions that may influence target and replacement behaviors, and (c) the development of systematic reinforcement and response systems in response to the demonstration of target and replacement behaviors (Horner, Sugai, Todd, & Lewis-Palmer, 1999; O’Neill et al., 1997). In this study, the results of the application of such plans were compared and quantified using a pretest and posttest measure that could be translated into an effect size. These results could then assist the researcher in investigating effective and efficient methods of conducting an FBA.
Behavior support plans (BSP) is considered a lower level of intensity intervention plan than a BIP (Horner, Sugai, Todd, & Lewis-Palmer, 1999). The level of prerequisite data is generally less than that of a BIP primarily consisting of indirect and direct descriptive methods of assessment. Similarly, a BSP addresses the instruction of alternative behaviors, environmental conditions, and reinforcement procedures.

Consequences are events that follow the occurrence of a behavior (Bijou et al., 1968). Further analysis of consequences may reveal events that maintain or reinforce the occurrence of a behavior. These events can be manipulated to help shape behaviors by positively or negatively reinforcing the demonstration of certain behaviors. This study examines the direct ABC and AB descriptive observational methods of Iwata et al. (1994) and Carr and Durand (1985) in which the former includes the analysis of consequences in the behavioral relationship.

Contingency is the relationship between behavior and the corresponding consequence (Skinner, 1953). Contingencies are important to this study in that the analysis of these relationships can suggest environmental and response changes that may increase the effectiveness of the intervention.

Direct descriptive observations are a method of FBA where a student is observed within a more natural as opposed to analog context to determine the frequency, temporality, intensity, or physical byproduct of a given behavior (O’Neill et al., 1997).

Function refers to the underlying purpose or communicative intent of a given behavior. Functions are generally to (a) get or obtain or (b) escape or avoid a wide-range of antecedents and consequences. This study assessed the effect of indirectly
inferred functions, directly observed functions, as compared with or combined with experimentally tested functions.

*Functional assessment (FA)* is a method of FBA that uses experimental design to test hypotheses based on the manipulation of the environment, antecedents, and consequences (Carr & Durand, 1985; Iwata et al., 1994). The FA is intended to establish a causal relationship between the behavior and the environment (Skinner, 1953).

*Functional behavioral assessment (FBA)* is a process for analyzing the antecedents, consequences, and function of a behavior or class of behaviors in the interest of determining a behavioral intervention (Gresham, 2001; Iwata et al., 1994; O’Neill et al., 1997).

*High-incidence disabilities* are considered, for this study, specific learning disabilities, emotional disturbance, other health impairment, and speech and language disorders. This study examined the effectiveness of FBA methods and intervention plans specific to the high-incidence disability population of school-aged students in the public-school system.

*Low-incidence disabilities* are visual impairment, deaf-blind, hearing impairment, orthopedic impairment, and traumatic brain injury. Unlike this current study, the majority of research related to the FBA process involved this population within a clinical setting.

*Positive behavioral supports* is a system founded on the principles of applied behavioral analysis that utilizes nonaversive procedures to (a) decrease the occurrence of negative behaviors, (b) increase the occurrence of positive behaviors, and (c) enhance lifestyles as a result of these learned changes (Carr et al., 1999).
Setting events are antecedent conditions that occur well before the behavior, yet are functionally related to the behavior, for example, missing breakfast before arriving at school (Bijou, Peterson, & Ault, 1968). Any examination of student behavior must consider setting events as they are difficult to control for and identify given the remote nature of the event and they can have a lasting and powerful effect on the student.

Typical intervention agent is an individual who is implementing a behavior plan and would have more natural contact with the recipient of the treatment such as a parent, teacher, school psychologist, instructional assistant, and so on (Carr et al., 1999). As this study did not address specifically the effects of typical versus atypical intervention agents, studies involving typical agents were included in this meta-analysis.

Typical intervention setting refers to a location where the intervention plan is implemented that is considered a more natural setting such as the individual’s home, school, community, and so on (Carr et al., 1999). This study is designed specifically to assess the effects of the assessment and implementation process within the typical setting of public schools. Studies with data that reflect multiple-settings, wherein behavioral change or lack thereof can be identified specifically within the public-school setting, were included in this meta-analysis. Those studies where the data were aggregated into all typical settings were not included in this study.

Summary

The IDEA and IDEIA mandate the use of FBA and positive behavioral interventions in addressing pervasive and maladaptive behaviors exhibited by students with disabilities. These same federal statutes support greater numbers of
students with high-incidence disabilities, such as specific learning disabilities and emotional disturbance, being included in the general-education environment at a higher level than in years past. Although there is a wealth of empirical data to support the varied models and methods of FBA, there are no clear guidelines for practitioners or agreement among researchers as to what constitutes an appropriate FBA. As a result, there is the research-to-practice gap. Limited resources, training, and ability to implement FBA and PBS within the public-school setting have compounded this difference between research and practical application. Therefore, this study utilized meta-analytic procedures to examine single-case studies addressing pervasive and maladaptive behaviors displayed by students with high-incidence disabilities attending public-school settings. The results of this meta-analysis may assist practitioners in developing procedures and templates for conducting a FBA and BIPs, which may save valuable time and resources and increase the likelihood of educators meeting the mandated assessment and intervention requirements.
CHAPTER II

REVIEW OF THE LITERATURE

The review of the literature provides the empirical foundation for frameworks that could be applied within the public school to facilitate the positive behavioral supports (PBS) approach and the functional-behavioral-assessment (FBA) process. There is emerging research that serves as a starting point for bridging the research-to-practice gap by involving educators as researchers and conducting assessments in typical settings with typical agents.

The purpose of this chapter is to define the overarching principles of PBS, FBA, and the various applications of empirically validated positive behavioral interventions. These principles discussed in the first section, serve as the context for the highly complex process of conducting a FBA and subsequently developing and implementing a behavioral intervention to decrease the occurrence of targeted behaviors and to replace them with appropriate alternatives. The second section includes a detailed exploration of the vast body of research related to the effectiveness of FBA procedures and intervention strategies (e.g., Ervin et al., 2001; Hanley, Iwata, & McCord, 2003; Stage & Quiroz, 1997). The synthesis studies assist the researcher in identifying factors that influence the assessment and intervention process by pulling from multiple studies. These generalizations can pinpoint areas of strength and limitations, which includes the utility, acceptability, and practicality of the FBA process.

Effectiveness of PBS and FBA

The first section of this literature review provides a philosophical context for the development and implementation of FBA-based interventions. PBS is the
synthesis and the application of previously established behaviorally-based practices that utilizes a systems approach to support parents, educators, and students. The efficacy of PBS as a framework for assessment and intervention at the district-wide, site-, and the individual-level is detailed in this section. The standard for effective PBS involves not only a reduction of problematic behavior but also the development of necessary replacement behaviors and skills that effect positive lifestyle changes in the social, educational, vocational, familial, and recreational arenas (Carr et al., 1999). A mainstay of PBS is the use of FBA to target behaviors. Numerous studies conducted since the late 1970s utilize the results of FBA-interventions along with quantitative research to evaluate the efficacy of PBS (Carr, 1977; Dunlap, Dunlap-Kern, Clarke, & Robbins, 1991; Ervin et al., 2001). The research around PBS consistently identifies the need to establish clear FBA guidelines that build upon strong internal validity through rigorous empirical research; however, these guidelines also must address the importance of social and external validity to increase the likelihood of generalizing to practitioners (Carr et al., 1999; Gresham, Quinn, & Restori, 1999; Hanley et al., 2003).

In an effort to evaluate the effectiveness of PBS and the factors that may modulate the effectiveness of PBS, Carr et al. (1999) conducted a review of the literature that included 216 articles from 36 journals. An initial hand-search was conducted of all relevant educational, medical, and psychological reviews. This review led to additional references through research articles, review papers, books, and newsletters that were screened by crossing disability diagnoses with behavior topographies using the following abstract index references: Child Development Abstracts and Bibliography, Current Contents/Social and Behavioral Sciences, ERIC,
MEDLINE, Psychological Abstracts, PsychINFO, PsychLIT, PsychSCAN/MR, and the Social Science Citation Index. The researchers also requested information from leading experts in the field as well as organizations that are involved in advocating for and providing services to individuals with special needs. The researchers developed their own index for measuring the effectiveness of the PBS quality standards as well as means to determine which studies to include and exclude from the review of the literature.

The team of researchers applied an inclusion and exclusion criteria for the selection process that included only peer-reviewed articles published in English between 1985 and 1996. Nonpeer reviewed publications were not included in order to establish a standard of experimental rigor. Studies not meeting the experimental rigor should have been included, coded, and analyzed for differences as recommended in the meta-analysis and research synthesis literature. In order for the study to be included in the review, the study participants had to have been diagnosed with a primary or secondary disability of mental retardation, autism, or pervasive developmental disorder (PDD). Participants with dual diagnoses were retained for the study. Only studies that addressed behaviors classified as aggressive, SIB, property destruction, and tantrums were included. For inclusion purposes, studies were classified as either antecedent-based or as consequence-based. Other types of interventions were excluded from the study. Of the original 216 single-case studies, only 109 studies with 230 participants were included in the final analysis. Of the 109 studies, 100 of the studies used single-subject experimental designs. The size ranges for the replicated design studies was 2 to 7 individuals with the average being 3 individuals in a study.
The researchers coded with respect to demographics, assessment practices, intervention strategies, and outcome measures. The following demographic variables were recorded: publication year of the article; gender; diagnosis or disability type; age; level of mental retardation, where appropriate; and the topography of the behavior. The categories of assessment methods were indirect, direct, functional analysis (FA), and multiple methods or any of the previous methods combined. Intervention category, systems change, and ecological validity were the three themes within the intervention strategies category. The intervention category was divided further into two generic categories stimulus-based and reinforcement-based interventions. Systems change was coded as either (a) a behavioral change on the part of a significant other, parent, teacher, employer, and so on or (b) a more global environmental change such as the restructuring of personnel, altering the living or work environment, and so on. Successful behavioral change or outcomes were measured as the percentile reduction in problem behavior. As the majority of the studies utilized a reversal design wherein the interval condition alternated with the baseline or vice versa, the reduction was calculated using the mean of the last three baseline data points less the mean of the last three intervention points. No justification was given as to why the last three points were used. The raters then coded intervention effects as the percentile of behavioral change.

Other outcomes measured included stimulus generalization, response generalization, maintenance, and social validity. With the exception of maintenance, the other measures were more subjective, and, as such, the method and rubric of measurement varied dramatically making the analysis of these measures difficult. Stimulus and response generalization involves the transfer of acquired skills and
behavioral change to other antecedents and setting events or consequences. The maintenance outcome refers to behavioral change in the form of a reduction of problematic behaviors and increased replacement behavior use over time. For this research synthesis, social validity was measured as feasibility, desirability, and perceived effectiveness.

The researchers estimated intervention outcomes through visual inspection measuring point by point to generate the last three baseline points and the last three intervention points. The intervention mean was subtracted by the baseline mean, divided by the baseline mean, multiplied by 100 to yield the percentile difference. Similar procedures were applied in evaluating the occurrence of positive behavior, stimulus generalization, response generalization, and maintenance. These outcomes were measured following the same percentile formula as above, however; the mean was calculated using the available intervention points, which in some cases were as few as one data point. These outcomes should be viewed with caution as fewer data points either can positively or negatively skew the results.

In many instances, an individual was subjected to more than one distinct type of treatment, in which case the outcomes were coded separately for each intervention outcome. When the same intervention, such as choice, was repeated through several phases of a reversal design only one outcome was calculated. Out of the total 230 individuals, 145 (63%) demonstrated a single outcome in contrast to 85 (37%) individuals who produced more than one outcome.

Reliability was calculated related to the inclusion and exclusion of articles, coding of continuous (i.e., positive behavior, problem behavior, generalization, maintenance, and lifestyle) and categorical (i.e., demographics, topography of the
behavior, etc.) variables, and data entry. Of the 109 included articles in the study and the 107 excluded articles, 50 articles from each group were rated by a coauthor of the study who had not been involved initially in the selection process. The coauthor applied the same inclusionary and exclusionary criteria to the selected articles. There was 100% agreement between the initial coder and the coauthor. A similar process was used for the coding of continuous and categorical data. One of the coauthors was given 7 randomly selected articles coded separately by each of the coders, which totaled 28 articles for the reliability sample. Continuous data reliability, based on the Pearson-product moment correlation, was +.99. The Kappa values for the categorical data ranged from .82 to 1.00, which represent a very high level of agreement. The keystroke error rate was only .11% between coders.

The researchers found a number of potential biases in the retrieved literature that met the rigorous inclusion criteria for the primarily single-subject experimental design studies. Particularly in the applied behavioral field, researchers strive to balance the rigor of experimental design with the relevance of practical application and generalizability. Experimental control dramatically increases the internal validity of a study. This increased internal validity can be at the cost of external validity, as this level of control generally requires atypical agents working in atypical settings where the venues of intervention are more restricted than in a typical setting like the individual’s home or school. In single-subject studies, external validity is compromised further when compared with larger group trials. This inclination toward more rigorous types of research is evidenced in the sources for articles included in the synthesis. The researchers suggested that these studies, not withstanding some limitations, do serve as a starting point for further experimental analysis as well as
review of nonresearch literature that may address ecological validity issues within
PBS and the FBA process.

The results of the study supported the overall effectiveness of PBS and the
FBA process as well as the presence of a research-to-practice gap. An atypical agent
such as a researcher or clinical psychologist was more likely to conduct the
assessments and interventions as compared with a typical agent such as a parent or
teacher. Over twice as many assessments and interventions occurred in atypical
environments such as a clinic or nonintegrated facility or school. It is noteworthy that
in all of the studies the typical provider acted under the guidance of an atypical agent
operating in a consultative role, although this is to be expected given the more
rigorously reviewed sources of the articles included in this study. The study also
analyzed assessment types, interventions, systems of change, contexts, generalization,
social validity, maintenance, lifestyle changes, and effectiveness of the interventions.

Over the decade leading up to the publication of the Carr et al. (1999) study,
there was a change in direction from single-method FBAs toward a multiple methods
approach. Of the 366 outcomes presented in the research, 200 were FA-based
interventions with 124 of those interventions developed by atypical agents. Moreover,
139 of the FA-based interventions were implemented in an atypical setting, and those
interventions were least likely to occur and subsequently generalize to all relevant
contexts (i.e., home, school, the community, etc.). In contrast, of the 57 non-FA or
Indirect- or Direct-descriptive-only methods, typical agents completed 41. Over half
of the non-FA interventions applied to all of the relevant contexts (e.g., home, school,
and community), which suggests a need to expand assessment and intervention
practices beyond clinical settings if the goal is to enhance quality of life. Given the
stake that most typical agents have in sustained behavioral change, typical agents were more likely to alter their behavior and the more removed atypical agent was less likely to modify their own behavior in interacting with a participant. The number of multiple method FBAs increased from 19 in the first 3 years of studies included in this synthesis to 162 of 266 studies by 1996, suggesting a trend in the research community toward more comprehensive FBA approaches. The trend is supported by what is found in practice currently in terms of more comprehensive FBA methods, however, the approach is no more clearly defined than in 1996 in terms of guidelines (Hanley et al., 2003).

The types of interventions implemented within the PBS framework have changed. Over time, stimulus-based interventions have gained favor perhaps in recognizing that it is often easier to manipulate antecedent and setting events through environmental changes, curricular adaptations, and changes in agent behaviors. Stimulus-based interventions assume a more proactive approach to addressing target behaviors. In contrast, reinforcement-based interventions have decreased, as they are more reactive in nature. Of the 266 studies, 102 used non-PBS interventions involving extinction, differential-response-to-other behavior (DRO), or punishment in the form of response cost consequences, brief restraint, or some aversive consequence.

In terms of the overall effectiveness of PBS interventions, 68% of the participants experienced substantial reductions of targeted behaviors as defined as a behavioral change of 80% or more from the baseline measures. Despite the trend toward more stimulus-based interventions, reinforcement-based interventions had substantial reductions 71.6% of the time as compared with 66.5% substantial
reductions for stimulus interventions. The percentage of outcomes for stimulus generalization demonstrated no trend, whereas response generalization increased slightly over time.

Illustrated in Table 3 is the relationship between assessment variables and outcome effectiveness, which is particularly relevant to the current study that seeks to establish guidelines for effective FBA methods and intervention practices. The data reflect that utilizing a method of FBA, any method, corresponded to a higher success rate of 59%. Direct observation, FA, and combined methods resulted in an approximately 60% rate of success. Indirect-only resulted in the lowest success rate of 42.1%; however, the low numbers of outcomes for both Indirect and Direct reflected should be viewed with caution.

Additionally, the socially mediated functions of behavior attention, tangible, escape, and multiple functions reflected a success rate of nearly 60%, whereas the nonsocially mediated success rate for sensory seeking was lower at 23.5%. This discrepancy may be due to sampling error as well. In assessing the validity and reliability of clinical methods as compared with more natural typical types of assessment, single assessments achieved a higher success rate of 60% than the repeated measures at 36.4%. These results may suggest that more time-consuming and intensive multiple assessments may not produce a higher rate of success than single assessments.

Concerning maintenance of behavioral reduction over time, PBS interventions, where documented, were relatively stable. Of the participants demonstrating a reduction of target behaviors of 90% or more over baseline, 68.7 to 71.4% maintained this level over a 2-year period following the initial treatment.
Studies conducted in atypical environments by atypical agents were less likely to contain a measure of maintenance than their typical counterparts.

Table 3
Relationship Between Assessment Variables and Outcome Effectiveness

<table>
<thead>
<tr>
<th>Factors</th>
<th>Outcomes</th>
<th>Successes</th>
<th>Successes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( f )</td>
<td>( f' )</td>
<td>( % )</td>
</tr>
<tr>
<td>Type of assessment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indirect assessment</td>
<td>19</td>
<td>8</td>
<td>42.1</td>
</tr>
<tr>
<td>Formal direct observation</td>
<td>10</td>
<td>6</td>
<td>60.0</td>
</tr>
<tr>
<td>Functional analysis</td>
<td>105</td>
<td>64</td>
<td>61.0</td>
</tr>
<tr>
<td>Combined assessment</td>
<td>132</td>
<td>79</td>
<td>59.8</td>
</tr>
<tr>
<td>Assessment conducted</td>
<td>266</td>
<td>157</td>
<td>59.0</td>
</tr>
<tr>
<td>No assessment conducted</td>
<td>100</td>
<td>32</td>
<td>32.0</td>
</tr>
<tr>
<td>Type of function</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attention</td>
<td>32</td>
<td>20</td>
<td>62.5</td>
</tr>
<tr>
<td>Escape</td>
<td>122</td>
<td>75</td>
<td>61.5</td>
</tr>
<tr>
<td>Tangible</td>
<td>27</td>
<td>15</td>
<td>55.6</td>
</tr>
<tr>
<td>Sensory</td>
<td>17</td>
<td>4</td>
<td>23.5</td>
</tr>
<tr>
<td>Multiple</td>
<td>52</td>
<td>33</td>
<td>63.5</td>
</tr>
<tr>
<td>Assessment repeated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>11</td>
<td>4</td>
<td>36.4</td>
</tr>
<tr>
<td>No</td>
<td>255</td>
<td>155</td>
<td>60.0</td>
</tr>
<tr>
<td>Assessment information used</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>231</td>
<td>135</td>
<td>58.4</td>
</tr>
<tr>
<td>No</td>
<td>35</td>
<td>22</td>
<td>62.9</td>
</tr>
</tbody>
</table>


Only 24 of the 109 studies examined quality of life or lifestyle change. Of those studies, only 2 demonstrated a 100% improvement as compared with baseline. Due to the relatively low numbers of typical agents involved in the studies, it is
reasonable to have a lower level of social validity measures. Furthermore, change in quality of life takes time and without typical agents to assist in the maintenance of such changes the likelihood of sustained lifestyle change appears to decrease given the results of this study. Desirability was assessed in 12 studies in which the interventions were viewed as preferable to highly-preferable as compared with previous methods of intervention. The metric for both acceptability and perceived effectiveness varied, which made these aspects of social validity difficult to assess; however, they were regarded favorably overall. The research encourages future investigation of both lifestyle change and social validity that includes the perspective of the participant.

Carr et al. (1999) provided compelling support for the effective and efficient use of FBA in effecting positive behavioral change and enhancing quality of life. Furthermore, they establish patterns of successful application across typical and atypical settings and agents. There is acknowledgement of the limited number of studies that involve typical and atypical settings and agents, which, given the complexity of the FBA process, brings into question the viability of FBA in public schools. The findings of the Carr et al. study, however, are broad in terms the use of multiple methods of assessment in developing these successful interventions and the social validity of the interventions. Although quality of life and positive behavioral change or the appropriate use of replacement behaviors are identified as intended outcomes of the positive behavioral intervention strategies approach, there were few studies that examined these goals. The emerging presence of these outcomes in the literature may be related to the limited involvement of typical agents who might have a greater investment and success rate in effecting these more long-term outcomes as
opposed to merely a reduction in behavioral excesses (Carr et al., 1999). As such, a later study by Kincaid, Knoster, Harrower, Shannon, and Bustamante (2002) addressed these areas that needed further investigation into the more positive and practical elements of PBS.

Kincaid et al. (2002) assessed positive behavior, quality of life, and social validity outcomes through FBA and positive behavioral intervention supports to address problematic behavior. Whereas Carr et al. (1991) used more outcome-based evaluations to assess the effectiveness of PBS, Kincaid et al. sought more person-referenced outcomes from the perspective of the individual implementing the intervention as well as the individual receiving the treatment.

To create a sufficient sample and address the areas of interest, the researchers contacted over 200 behavioral intervention teams trained in the implementation and technical support of PBS with individuals between the ages of 3 to 22 years of age in the Tri-State Consortium in Pennsylvania, West Virginia, and Virginia. Nearly all of the agents and settings were considered typical (i.e., parents and educators working in schools or the home). Each member of the team, including the individual student or adult client, received a 3-point Likert Scale Behavior Outcomes Survey (BOS) and a 5-point Likert Scale Quality of Life Survey (QLS) or interview. The BOS accounted for behavioral occurrence as more or less frequent, more or less severe, and shorter or longer periods. The QLS measured change in the following areas: interpersonal relationships, self-determination, social inclusion, personal development, and emotional well-being. Of the more than 200 teams, 78 teams consisting of 379 individuals completed both surveys.
The researchers originally intended to provide the teams with several surveys over the span of at least a year. In this manner, the researchers would conduct more of a longitudinal analysis related to the outcomes. It does not appear from the literature that more than a few teams completed more than one survey. The surveys that were completed were all done postintervention.

The researchers compiled the data from the surveys and interviews, and means were calculated for the focus areas of intervention efficiency, alternative skills development, efficiency of teaching alternative skills, and quality of life. The results of the surveys were cross-referenced using direct observation methods by the researchers; however, these results were not included in the study and there was no mention of the level agreement. An analysis of variance (ANOVA) was used to examine for any differences between the respondents for the subscales of the QLS.

In terms of behavioral change, teams provided their perceptions about the occurrence, severity, duration, and the development of alternative skills or replacement behaviors after the implementation of PBS interventions. Eighty-two percent of the respondents indicated that target behaviors occurred less frequently following PBS training. Similarly, 78% of the respondents identified a lower level of severity, and 76% indicated that the duration of the behavioral excesses decreased.

One of the primary differences between PBS and more classic ABA interventions is the emphasis on teaching replacement behaviors. This shift in the literature toward alternative-skills development is reflected in the increased number of respondents (71%) who indicated that these replacement behaviors occurred more frequently following PBS intervention as opposed to other traditional methods. Eighty-eight percent of respondents reported that these appropriate behaviors occurred more
frequently, and 76% indicated a more independent use of these replacement behaviors. These positive perceptions toward the effectiveness of PBS interventions as measured by decreases in target behaviors and increases in the use of replacement behaviors also were consistent with attitudes toward the efficiency of PBS.

Social validity or perceived efficiency was divided into several general or variations of questions that were presented through the surveys: *How well did it work?* *How comfortable were you in implementing the interventions and strategies?* *How consistent was the implementation?* and *What was the extent of the interference of these interventions?* An acceptable level of efficiency was considered a score of 4 to 5 on a 5-point scale. Sixty-seven percent of the respondents rated the effectiveness of the interventions as working well. Over 91% of the respondents expressed that they were very comfortable with the support strategies and teaching replacement behaviors. Additionally, 81% of the respondents reported consistent implementation, and 78% indicated little or no interference with daily routines. These results suggest an increased level of social acceptability among practitioners as well as efforts to bridge the research-to-practice gap.

The results of the Quality of Life surveys presented only modest overall gains ($M=3.79$) attributed to PBS interventions. The results subscales of QLS appeared as follows: interpersonal relationships ($M=3.79$), self-determination ($M=3.62$), social inclusion ($M=3.96$), personal development ($M=3.35$), and emotional well-being ($M=3.27$). These scores do not necessarily reflect change over a predetermined period of time but rather at a given point in time, which may account for the slight gains in quality of life.
With reference to the perception of quality of life across agents, the results of the ANOVA reflected consistently higher ratings overall quality of life ratings by parents and direct-care providers for overall quality of life, interpersonal skills, self-determination, and social interaction. School administrators reported consistently lower levels for overall quality of life and other subscales as compared with other respondents. Researchers suggested that the difference reported between administrators and other respondents likely was related to the amount of direct contact with the individual. Direct-care providers and parents were in a better position to observe substantive life changes over a period.

In conclusion, studies by both Carr et al. (1999) and Kincaid et al. (2002) found PBS and FBA methods to be effective methods for reducing the occurrence of targeted behaviors and increasing the use of replacement behaviors. Both studies suggested that the social validity for FBA and positive behavioral supports by typical agents within typical settings increased in terms of desirability, acceptability, and perceived effectiveness. Furthermore, these studies urged the research community and practitioners alike to pursue measures relevant in applying these empirically validated methods of addressing pervasive and maladaptive behaviors. Because PBS is a systems approach for effecting positive behavioral change, additional research is needed to assess its practical application at a district as well as site levels in order increase the social validity and generalizability of PBS and the FBA process within public-school settings.
Synthesis, descriptive, and meta-analytic studies related to FBA methods and positive behavioral interventions

This second section examines effectiveness of various types of behavioral interventions across environments and agents before further refining the analysis of FBA procedures. A number of literature synthesis studies, descriptive analyses, and meta-analyses are reviewed in this section and illustrate factors that modulate the effectiveness of the FBA process from inquiry to implementation. All of the studies, at least, address pervasive and maladaptive behaviors. The majority of the studies reviewed involve atypical agents (i.e., researchers, clinical psychologists, and graduate students) in atypical locations (i.e., clinics, residential facilities, nonintegrated school settings, etc.). Two of the studies included in this review focus solely on the development of a FBA and interventions in the public-school setting, although only one included a practitioner as the agent. Notwithstanding the atypicality of the agents and locations and the experimental nature, the results provide valuable insight into the effective implementation of FBA-based interventions.

Functional behavioral assessment is an effective process for analyzing behavior through systematic observation, assessment, and hypothesis testing to develop a BIP (Hanley et al., 2003; Horner, 2004; Horner & Sugai, 1999; O’Neill et al., 1997). Although there is a large body of research dedicated to the development of FBA and BIP, the majority of these studies do not take into account students with more commonly occurring disabilities within the public-school system such as specific learning disabilities, other health impairment, emotional disturbance, and speech and language disorders (Bergstrom, 2003; Dahlstrom, 2003; Sugai, Lewis-Palmer, & Hagan-Burke, 1999). The dilemma for practitioners in conducting an FBA
and developing a BIP is twofold given (a) the emphasis in the federal mandates to include students in the mainstream environment without procedural guidance in the process and (b) the need to bridge the research-to-practice gap (Horner, 2004; Sugai, Horner, et al., 1999).

The IDEA (1997) and IDEIA (2004) amendments require the application of FBA and BIP within the public-school setting to address pervasive and maladaptive behavior. These same mandates require that students be educated to the greatest extent appropriate within the mainstream environment. As such, the number of students with more commonly occurring disabilities is increasing (OSEP, 2000). To address the increased need within school settings, systems such as PBS are implemented within schools to maximize the efficiency and effectiveness of already limited resources (Carr et al., 1999; Koegel, Koegel, & Dunlap 1996). A multitiered PBS approach provides a strong foundation of behavioral and academic success that enables staff to target students who are at-risk or in need of FBA and intensive intervention (Borgemeier, 2005).

Although FBA has been an evolving technology that is derived from the seminal works of B.F. Skinner and the founders of ABA principles, the research community has only recently begun to evaluate specific methods of assessing pervasive and maladaptive behavior (Dahlstrom, 2003; Hanley et al., 2003; Horner, 2004). These 3 methods of assessment, which include indirect, direct descriptive, and FA, can be used independently or in conjunction with other methods. Within each method are multiple variations, instruments, and approaches that include the FACTS, FAI, ABC, AB, and archival reviews to name a few. The process, however, can be extremely complex and requires a foundation in behavioral theory and applied
practice, which makes conducting a FBA and developing a BIP difficult at best within the public-school setting (Horner & Sugai, 1999; Nelson, Roberts, Mathur, & Rutherford, 1999).

Given the plethora of studies related to PBS, FBAs, and BIPs and the interest among the research community and practitioners to establish effective and efficient guidelines for assessments and interventions, it may be beneficial to review descriptive, synthesis, and meta-analytic studies that pool studies to discuss trends and patterns as recommended by Carr et al. (1999). The majority of synthesis research related to the FBA process is focused either on the outcomes of interventions or on the psychometric qualities of the individual methods. Other studies examine the process from the assessment to the implementation and in some cases beyond to maintenance of the intervention (e.g., Carr et al., 1999; Hanley et al., 2003). This section of literature review began with a meta-analysis of interventions intended to reduce problematic behaviors within the public-school setting. Assuming that interventions can be implemented successfully within the school setting and that there is an apparent relationship between the behavior and student-related variables, it is reasonable to analyze the effectiveness of the FBA process and the interventions across a variety of agents and settings through a meta-analysis. Therefore, the current study specifically targets FBAs and BIPs within the public-school setting, the focus of the review narrows to examine assessment and intervention procedures in this underrepresented context in the literature.

*Interventions to decrease problematic behavior in schools.* Disruptive behaviors within the school setting are indicative of lower levels of academic and social engagement, lower inclass performance, and poor performance on standardized
assessments and can be an early indicator of future antisocial behavior (Shinn, Ramset, Walker, Steiber, & O’Neill, 1987). From teacher training to ongoing staff development programs, the educational community is ill equipped to meet the needs of students who demonstrate pervasive and maladaptive behaviors (Horner & Sugai, 1999; Kaufman & Wong, 1991). Addressing these behaviors has become a primary focus of recent educational movements, such as PBS (Carr et al., 1999). A large number of school-based intervention programs have been evaluated. There are, however, few studies that have pooled the results of these studies to assist practitioners in being able to identify effective interventions that correspond to the topography of the target behavior within a given setting.

Stage and Quiroz (1997) conducted a meta-analysis of 99 studies that used interventions to decrease the occurrence of disruptive behavior in the public-school setting. The researchers used PsycINFO to search for studies from over 20 journals that yielded 310 abstracts. Using one source for a literature search, however, is not a comprehensive literature search or one that would yield studies from a variety of literature types of source, which may result in a bias. Of these abstracts, 273 studies were collected. There were 5,057 participants in these studies. Of those students, 4,117 were regular-education students, 55 were diagnosed with Attention Deficit Disorder Hyperactive type (ADHD), 57 were identified as having Emotional Disturbance (ED), 81 were identified as having a Learning Disability (LD), 31 were eligible for special-education services under the category of Mental Retardation (MR), 56 were identified as having multiple disabilities, 5 were with a Hearing Impairment (HI), and 550 were identified within the clinically significant range for
aggressive behavior or as having Oppositional Defiant (ODD) or Conduct Disorder (CD).

The interventions assessed were either behavioral, cognitive-behavioral, individual counseling, parent training, or multimodal. Behavioral interventions included timeout, overcorrection, token economies, group- and home-contingencies, self-management or monitoring, DRO, differential-reinforcement of low-rate responding (DRL), planned ignoring, response-cost, extinction, as well as function-based consequences. Contrary to PBS principles, the researchers inappropriately indicated that behavioral interventions do not involve the teaching of replacement behaviors. It appears that behavioral instruction of replacement behaviors is part of the cognitive-behavioral construct with treatments such as relaxation and social problem-solving. The individual counseling intervention is aligned with traditional psychotherapy. Parent training consisted of educating parents to serve as the agent for implementing differentiated reinforcement, compliance training, and timeout. The multimodal intervention combined two or more of the aforementioned interventions.

Four criteria were used for inclusion in the meta-analysis. First, studies had to provide a valid quantitative measure of the disruptive behavior at baseline and the intervention outcome. Studies where the dependent measure was deemed too subjective (e.g., teacher referrals to the office) or too global (e.g., anxiety or depression rating scales that do not capture disruptive behavior specifically) were excluded. Second, the studies had to be conducted in public-school settings. Third, studies that did not provide statistical results or included nonparametric tests, multiple analysis of variance, or analysis of covariance were excluded to minimize the confounding the recovery of effect size as a result of controlling for other variables.
In addition, only studies with 10 data points in at least one of the baseline or treatment phases were included to increase the accuracy of the ITSACORR program used in evaluating difference between the baseline and treatment phases. Fourth, the study needed to provide a nontreatment control group for comparison.

Of the studies included in the meta-analysis, the majority of the studies employed an $A_1B_1A_2B_2$ design with $A$ representing baseline conditions and $B$ the treatment conditions. The researchers developed their own unorthodox method of calculating effects size that is not found in other meta-analysis or literature reviewed in this study (cf., Busk & Serlin, 1992). The data points were reconfigured to $A_1A_2B_1B_2$ so that the data points represented a single baseline condition and treatment condition. Similar procedures were applied to studies with multiple baseline and treatment designs. The researchers entered and analyzed the data using the ITSACORR program to assess for threats to internal validity such as history or maturation. Because there were no statistically significant threats to the validity of the results, the effect size between experimental and control groups was calculated using $\Delta_{E-C} = \frac{M_E - M_C}{SD}$ in which $E$ represents the treatment group, $C$ equals the control group, $M$ is the mean, and $SD$ is the pooled variance between the treatment and the control group. This calculation may be flawed, however, in that there is an assumption that the variances are equal between the groups and that the treatment phase was not different from the baseline phase, which may not be correct.

Effect size for the interventions was calculated as the reduction in the target behavior using the mean difference between the baseline and treatment data phases divided by the pooled standard deviation of the two phases. The statistical significance between groups was determined by one-way ANOVA followed by
inspection of all differences between pairs of means by Student-Newman-Keuls test. According to the researchers, the assumptions for homogeneity of the effect sizes were met; however, there is little elaboration. Therefore, these results should be viewed with caution. Due to the limited number of studies for some interventions, such as punishment and parent training, the effect sizes were not considered reliable estimates but were included nonetheless. A total of 223 effect sizes produced a statistically significant mean effect of -.78 and a standard deviation of .58, suggesting that on average there was a behavioral reduction across all treatments among 78% of the treated students. The results of the \( t \) test showed that group contingencies (ES=-1.02), self-management strategies (ES=-1.00), and differential reinforcement techniques (ES=-.95) were, on average, more effective than functional assessment (ES=-.51) and cognitive behavioral strategies (ES=-.50). The relatively lower level of effectiveness may be related to the researcher’s misinterpretation of FBA procedures that is limited to consequence-based interventions that do not involve the teaching and reinforcement of replacement behaviors. The mean effect sizes for consequence types (i.e., negative, positive, combined, or none) were not statistically significant.

The study also examined effect size by educational category and clinical population as well as by grade level and setting. There was a statistically and practically significant difference in response to treatment between students with ED (ES=-.98, SD=.75) compared with students with ODD or CD (ES=-.48, SD=.27), suggesting that students with ED respond more favorably to interventions than students with ODD or CD. This is a reasonable assertion given that the fundamental distinction between the two categories is the volitional nature of the behaviors demonstrated by students with CD as compared with the limited locus of control in
students with ED. In other words, CD related behaviors generally involve a higher degree of motivation and choice than ED. The age of the participant was not a statistically significant factor; however, educational setting was. Students demonstrated greater reductions of behavioral excesses in self-contained environments compared with students in regular-education settings. The study did not address the issue of the agent’s role in assessing the behavior, implementing the intervention, or conducting the research.

A more recent meta-analysis conducted by Wilson, Lipsey, and Derzon (2003) assessed the effects of school-based intervention programs on aggressive behaviors. The primary source of studies was a search of bibliographic databases, including PsycINFO, Dissertation Abstracts International (DAI), ERIC, Medline, and U.S. Government Printing Office Publications. The limited number of sources used in the primary search may have narrowed the breadth of the analysis and introduced bias into the sampling. Meta-analyses, literature reviews, and the bibliographies of retrieved studies were reviewed further for additional references related to school-based intervention programs. Studies were selected based on the following criteria: (a) the study was reported in English, no earlier than 1950, and involved school-aged students (preschool through 12th grade); (b) the study assessed intervention effects for at least one outcome variable that represented broadly defined aggressive behavior; and (c) the study was experimental, quasi-experimental, one-group, and multiple-group design where at least one qualifying outcome variable was measured before and after the intervention.

Trained research assistants familiar with social-science research coded the included studies. Characteristics of the agent, setting, intervention type, and
participant were coded to assist in evaluating the interaction between the relevant factors and behavioral change. The specific effect-size statistic for this study was defined as the difference between the posttest mean and the pretest mean for a single-subject group divided by the pooled standard deviation of the pre- and posttest values (Wilson, Lipsey, & Derzon, 2003) or

\[
d_{ij} = \frac{Y_{ij} - X_{ij}}{s_{ij}}
\]

where \(X_{ij}\) is the group pretest mean, \(Y_{ij}\) is the posttest mean, and \(s_{ij}\) is the pooled standard deviation for Group \((j)\) in Study \((i)\). In this meta-analysis, the terms pretest and posttest appear to refer, respectively, to the baseline and treatment phases of the included studies. None of the reviewed studies appear to include pre- and posttest measures but rather measurements of pre- and postintervention conditions. Many case studies use multiple groups within a given study, which is why it is important to differentiate between group and study. Standardized means that the change effect size is weighted by the inverse of the sampling error variance of the mean, which appears as follows:

\[
\text{var}(d_{ij}) = \frac{2(1-r_{ij})}{n_{ij}} + \frac{(d_{ij})^2}{2n_{ij}}
\]

The \(n\) refers to the size of the sample of a given group \((j)\) within the study \((i)\). The correlations between pre- and posttest \(r\) are necessary data to compute the appropriate weights for the studies. These correlations were obtained from the primary study. When the correlations were not available from the primary study, the correlation was estimated from studies with similar dependent measures or from test-retest reliability coefficients from the primary study.
Two-hundred and twenty-one studies were included in the meta-analysis yielding 522 subject pre-posttest effect sizes. These effect sizes were analyzed in terms of a number of variables: research group and design, age, risk level, specific intervention type, and routine intervention type. The specific intervention types included social competence with no cognitive-behavioral component, social competence with a cognitive-behavioral component, behavioral or classroom management, therapy or counseling, multimodal (e.g., a combination of peer mediation, counseling, parent training, academic intervention, and the like), academic services, schools within schools, and peer mediation.

Tests of the homogeneity of the pre- and posttest effect sizes using the $Q$ statistic (Hedges, 1981) showed a higher level of variability across the intervention and control groups than expected given sampling error. In other words, the intervention group ($Q(333)=2,917$) and control group ($Q(187)= 924$) included studies that produced, in some cases, effect sizes that were either disproportionately larger or smaller than the corresponding mean across studies. Given this variability, the researchers would only be able to compare behavioral change between the intervention and the control groups; questionably, the study goes on to identify the overall effects for the various interventions.

In terms of the research design and group, there was a statistically significant difference of effect size across all of the intervention groups compared with the control group. The mean pre-posttest change for the control group was nearly zero suggesting that on average there is no change in aggressive behavior without intervention. There was no statistically significant difference between randomized (ES=.32), nonrandomized (ES=.16), and one-group (ES=.23) designs. The results
revealed a curvilinear relationship between age and effect size in that preschool and high-school students demonstrated the greatest behavioral change between control and intervention, respectively (ES=.27; ES=.30).

These researchers also examined the effectiveness of school-based intervention programs. The risk variables or the intervention groups, all demonstrated statistically significant differences in behavioral changes between the treatment and control groups. The effect-size differences were as follows: targeted (ES=.36), selected (ES=.26), and general population (ES=.09). These findings reinforce the notion that PBS is an effective approach for providing targeted interventions that are beneficial to all students but more importantly have a more profound impact on students at the highest level of risk.

Because of the limited number of studies (26) that reflected interventions developed and implemented by typical agents, the researchers analyzed the effectiveness of typical and atypical agents separately. Within atypical programs, or the 126 demonstration projects as they are referred to in this study, behavioral programs (ES=.43), counseling (ES=.41), schools-within-schools (ES=.32), and academic intervention (ES=.28) intervention programs had the strongest effect on behavioral change. Whereas the studies involving atypical agents provide insight into the efficacy of the various intervention methods within the public-school setting, the relatively small number of studies conducted by typical agents yielded an overall effect size of .10. Only the academic programs demonstrated a positive change in behavior; however, the difference was not statistically significant. The results of this comparison illustrate a ratio of nearly five atypical to every one typical study published. This review of the literature suggests clinically based studies are more
effective. The difference in not only the quantity of the studies published using typical agents but also the quality of those few studies further illustrates the need to bridge the research-to-practice gap.

In sum, the meta-analyses support the effective use of interventions to address challenging behavior at school. These primarily experimental design interventions also appeared to increase the reduction of problematic behaviors in correspondence to the severity of the participant’s needs, that is, students with higher risk levels and ED responded to a more positive degree, whereas students with lower levels of need or less severe disabilities responded to a lesser degree. Although behaviorally based interventions consistently demonstrated a high level of effectiveness, the complexity of the related assessment and limited experience may be barriers for generalizing these practices to practitioners. The dilemma for researchers and practitioners is how to conduct the assessments and develop interventions in a manner that is effective, efficient, and acceptable for general practice within the public-school setting. These meta-analyses did not address specifically the multiple methods of FBA, only interventions that may or may not have been derived from an appropriate assessment. Without practical guidelines for conducting a FBA and implementing a BIP, navigating the large number of studies to determine appropriate methods and corresponding interventions may be confusing for many practitioners. To that end, the next section of this literature review examines studies that refined the assessment process through the indirect, direct descriptive, and FA methods.

*FBA methods.* The functional analysis methods by Iwata et al. (1994) and Carr and Durand (1985) began the evolutionary process of conducting a comprehensive FBA. These methods involved the application of empirically based assessment,
control techniques, and interventions. The search for FBA methods included a search of *Current Contents, PsycINFO*, and *ERIC*, which included publications and unpublished sources. Dissertation abstracts were not considered and could result in bias. The abstracts from the databases were reviewed further for additional sources. In an effort to synthesize current FBA literature, Hanley et al. (2003) reviewed 790 published works before including 277 empirical studies that met the inclusionary criteria. Studies included in this quantitative review made use of (a) a preassessment to narrow variables prior to treatment, (b) a direct observation or measurement across at least two environments, and (c) some environmental manipulation of variables as means to assess the relationship between the environment and the behavior. Problem behavior in this study was defined as socially significant behavioral excess, such that it elicits a negative response from a reliable reporter in that environment. These behaviors may affect the acquisition of new skills by the individual or others in that environment, be considered harmful or dangerous, or require a more restrictive placement or living arrangement.

The review of literature included analysis of behavioral change between the baseline and intervention conditions through visual inspection of the results of the single-case studies. Furthermore, characteristics were coded by the researchers and trained research assistants familiar with ABA research. The coded methodological characteristics included model type (i.e., ABC, AB, or both), supplementary assessments (i.e., indirect, direct descriptive, or both), reinforcement type (i.e., social negative, social positive, or automatic), number of test conditions, assessment length, session duration, experimental design, and data presentation.
FA model types included antecedent-behavior (AB; Carr & Durand, 1985) and antecedent-behavior-consequence (ABC; Iwata et al., 1994) types. Specifically, 201 studies used the ABC model were compared with 56 AB models. Twenty studies used both the AB and ABC, which likely were employed in a two-level FBA design (see Figure 3 in Chapter 1). None of the literature utilized in this review included a reinforcement or a consequence behavior model as found in Carr et al. (1999), suggesting a more recent trend away from this model of FA. Of the FA methods examined, 13% included brief FAs compared with 87% of the studies that included full FAs. Given the preference for more rigorous FA methods, few studies employed Indirect (12 studies) and Direct descriptive (23 studies) as independent methods of FBA. Only seven studies included Indirect, Direct, and FA methods in conducting a comprehensive FBA.

Functional analysis methods follow a single-case study design in that the control and test conditions are compared in order to develop an effective intervention. The test condition involves some independent variable that may influence behavior, whereas the control condition generally reflects the absence of the independent variable being assessed. Most FA studies assessed both social-negative (89.2%) in the form of escape or avoidance and social-positive (82.5%) in the form of tangible and attention types of reinforcement. Given the likelihood of multiple element conditions that occur frequently in natural settings, 89.5% of the FA studies assessed the influence of negative versus positive reinforcement or social versus automatic reinforcement. This refined evaluation process for controlling and competing relationships among variables increases the effectiveness of interventions through
more precise function matching and minimizes program changes that may not influence the occurrence of the desired replacement behavior.

The majority of the studies utilized multi-element design (81.5%) in which the effects of several independent variables could be assessed over short intervals of alternating conditions. Nearly 16% of the studies utilized a reversal of ABAB design in which a single variable was introduced, withdrawn, and presented again. This design was most prevalent among AB studies or studies in which a single variable was being evaluated.

Of the 277 studies included in this review, 536 individual data sets were graphed with at least one data point documented per observation. The results of these data sets suggested differentiated outcomes or reductions of behavior in 96% of the cases as determined through visual inspection. Social-positive reinforcement accounted for 35% of behavioral maintenance as compared with social-negative (34%). Automatic reinforcement was relevant in 16% of the studies. Finally, reinforcement with multiple variables applied in 15% of the cases. With only 4% of the studies reporting an undifferentiated outcome, it is reasonable to assume that FBA methods are effective tools in assessing problematic behaviors and identifying maintaining functions.

Although the Hanley et al. (2003) study offers compelling support for the utility of FBA, specifically FA methods in determining the function of behavior, the research neither compares the various FBA methods nor provides empirical data to expand on the social validity of the FBA process by analyzing efficacy in typical environments. The study does present a very technical and meaningful analysis of the practical differences between the AB and ABC methods. The relative effectiveness of
the various methods of FBA and generalizability to typical settings remains for future studies.

*Descriptive analysis and critique of literature on school-based FBAs.* The descriptive analysis and critique of available FA research by Ervin et al. (2001) was the only study of this type that addressed the gap between the research community and practitioners by examining school-based assessments. This descriptive analysis and critique included 100 articles published between 1980 to 1999. This study was intended as a point of departure in establishing a framework for the practical application of FA and FBA procedures by examining prevalent features most applicable to practitioners. This study differentiated *functional assessment* as a range of procedures or methods, whereas *functional analysis* was the systematic manipulation of the environment to test an operationally defined hypothesis. This study also introduced the multiphase model of conducting a FBA (see Figure 2 in Chapter 1). Studies included in this analysis had to include at least one of the four phases detailed in Figure 2, which included any data collection to form a testable hypothesis and maintaining variables that influenced behavior. The methods of FBA could include Indirect, Direct, Analogue procedure, or FA. The research had to appear in a journal and the intervention had to be conducted in a school setting.

The researchers began their search for possible articles using the *PsychLit* and *ERIC* databases. A bias is made when only published material is reviewed. Unpublished studies should have been included and results compared those of published articles. The use of these sources did limit the scope of the search. To expand the search, potential articles were scanned further for additional sources. Another search of the same databases was conducted using initially screened authors
and journals that might contain relevant material. One of the authors scored all articles included in the analysis. The researchers coded relevant variables such as student demographics, setting, context (i.e., natural, analog, or both), behaviors targeted (i.e., appropriate and nonappropriate), FA type, assessment measures, manipulation of variables (i.e., type of variable such as antecedent or consequence, who did the manipulating, and experimental design), intervention type, and agent implementing the intervention. A review of the visual display was conducted to determine the intervention outcome where raters indicated whether the intervention was “effective,” “not effective,” or “unclear.” There was no statistical analysis of the results of the interventions.

The 100 studies included a total 238 students. The majority of the participants in the studies were diagnosed with one (35%) or more (54%) disabilities. Seventy-one percent of the students presented with a level of MR. The study did not differentiate clearly whether the remainder of the participants had either concomitant or independent speech and language disorders (SLD; 45%), physical impairments (25%), emotional or behavioral disorders (EBD; 18%), or LD (4%). It appeared that only 31% of the participants were identified educationally as having an independent high-incidence disability of SLD, EBD, or LD. This population is relevant to the current study that seeks to identify assessment and intervention guidelines for students with high-incidence disabilities. Over 52% of the FBAs occurred within the special-education setting. Other settings for the assessment included special education combined with home or clinic (12%), general education (12%), nonintegrated school (10%), or a research clinic (8%). Students not identified as having special needs were likely to have the FBA conducted within the general-education setting. The
assessment process occurred in an analog-context 43% of the time compared with
36% of cases in a natural-only setting for all students. Disruptive behaviors, such as
screaming and throwing an object (48%), aggression (41%), SIB (30%), off-task
behavior (24%), and property destruction (12%), were most common among
participants. The multiphase assessment process provided some of the most relevant
results in establishing an effective framework for evaluation.

The multiphase approach was divided into three areas for analysis: descriptive
phase, experimental analysis or hypothesis testing, and intervention procedures. Most
of the agents conducted a descriptive assessment phase (74%), whereas over 90%
conducted an experimental analysis of the target behavior. Just over half of the
participants received an intervention derived from the assessment phase or phases.
The interpretive phase was not included in the analysis most likely because of the
difficulty in quantifying the analytic process involved in developing the hypothesis.
Systematic observations were used with all but two of the participants and only 42%
of the agents employed descriptive observations, they were always done in
conjunction with systematic observations. Interviews were incorporated into the
assessment process in 49% of the cases.

The experimental analysis phase included the type of variable manipulated
and the agent manipulating the variable. During this phase of assessment, 69% of the
studies reflected manipulations of both antecedents and consequences, or commonly
the ABC model. Antecedent-only or AB manipulations accounted for 12% of the
studies in contrast to reinforcement- or consequence-based manipulations, which
were included in only 9% of the studies. The study also noted that students without
disabilities were more likely to receive consequence manipulations, whereas students
with disabilities were more likely to experience antecedent manipulations. Predictably, the experimenter working independently (53%) was most likely to manipulate variables, and the educator working with the experimenter (14%) was least likely. In terms of implementation, school personnel were most likely to be the agent and did so without assistance in 23% of cases. Experimenters implemented the intervention 10% of time and were less likely to provide assistance in special-education settings as compared with general-education classrooms. These results appear to support the assumptions that ABC is emerging as the method of preference among assessors and that a reactive, consequence-based approach is gradually phasing out. The analysis related to the individual may indicate that there is a difference in levels of confidence in researchers and practitioners in manipulating variables.

The data reviewed in this study were consistent with other synthesis studies in acknowledging that treatment acceptability and social-validity are absent generally from the literature. Only 12% of the studies reported treatment acceptability, which was considered acceptable to highly acceptable by agents. Social validity outcomes were only reported in five cases. This information was documented through informal ratings or direct observation, and the outcomes were reportedly favorably. As with nearly all of the previous synthesis research in the behavioral field, acceptability and social validity needs to be studied further, particularly if the goal, as stated in the Ervin et al. (2001) study, is to establish a framework for an effective and efficient FBA process.

This descriptive analysis supports the utility of FBA procedures; however, there is limited guidance in terms of the multiphase approach, suggesting further
The existing research for application in the public-school setting is limited across settings, agents, behaviors, and types of interventions. The variation in methods and procedures across and within studies also makes it difficult to establish generalities even with pooled data. Furthermore, certain practical and ethical considerations inherent to expert consultation and experimental design make application in typical settings problematic. Social validity and acceptability are ignored within the literature, which may limit the feasibility of practical application. In other words, FBA is an effective method of altering problematic behavior, however, implementation and generalization to the school setting requires further study.

The utility, acceptability, and practicality of FBA. An analysis of 97 FBA studies reviewed found that 88% of the participants had severe, low-incidence disabilities, 61% of those studies were conducted in a clinical setting, 42% of the target behaviors were self-injurious, and 100% of the assessors and implementers were not practitioners (Nelson et al., 1999). In light of these findings, there was a clear need to develop procedurally- and empirically-sound guidelines for practitioners that are effective and efficient given the current context of the public-school system (Broussard & Northrup, 1995; Horner, 2004). As a follow-up to the earlier Nelson et al. (1999) study, Reid and Nelson (2002) conducted a descriptive analysis to examine the utility, acceptability, and practicality of FBA for students with high-incidence disabilities. Of particular relevance to this current meta-analysis, the Reid and Nelson (2002) descriptive analysis was the only study to examine the social validity of FBA in typical settings with high-incidence disabilities.
A computer search of the *Exceptional Child Education Resources Abstracts* and *Psychological Abstracts* was conducted to identify articles related to behavioral interventions with students with high-incidence disabilities. Although the initial computer search was somewhat limited in number of sources, a hand search of the *Journal of Applied Behavior Analysis, Journal of Experimental Analysis of Behavior, Behavioral Disorders, and Journal of Behavioral and Emotional Disorders* was completed as well to identify related studies. As with the previous study, only published articles were included resulting in bias as many unpublished studies exist and could provide relevant details for the follow-up study by Reid and Nelson (2002). Studies were included in the research if a FBA was conducted for students with high-incidence disabilities within a public-school setting. Fourteen studies met the inclusion criteria, which included 43 participants. Thirty-seven of the participants were boys, and six were girls. Eligibility and diagnostic information was available for 17 participants where 7 were diagnosed with ADHD and 10 with EBD. Twenty-five were studies conducted in general-education classrooms; the remainder occurred in a self-contained special-education classroom (10) and special schools (8), which were assumed to be nonintegrated public schools.

The measures for determining the success of an intervention, acceptability, and practicality are unclear as the review of the literature relies on qualitative reporting instead of any statistical analysis or discussion of a visual inspection of the results. The researchers identified five questions to assess the extent to which the research community had begun to address the utility, practicality, and acceptability of FBA procedures: *Did the FBA procedures used by researchers improve behaviors of students in school settings?*, *Do direct service providers perceive the treatment effects
associated with FBA as socially significant?, Do direct service providers perceive the
treatment effects associated with FBA as socially acceptable?, Have school personnel
performed the FBA?, and How demanding is performing the FBA in terms of time and
performance?

In response to the utility of the FBA process, there was a favorable behavioral
change, as defined as at least an increase in appropriate replacement behaviors or a
decrease in target behaviors, in 12 of the 14 studies. Only two studies reflected slight
behavioral improvements; however, these behavioral changes were not defined
quantitatively. Researchers in seven studies reported a behavioral reduction from
baseline to nearly zero and an increase in appropriate replacement behaviors to nearly
100%.

The generalizability and durability of these interventions was not ascertained
easily as only one study examined maintenance of the treatment effects. Given the
functional relationship between academic tasks and target behaviors in students with
high-incidence disabilities, the intervention often consisted of simple curricular
adaptations, such as choice of assignments or order of task completion, moderating
task difficulty, and providing additional prompts. This relationship between academic
tasks and behavior may enhance acceptability of the FBA process with practitioners.

Social acceptability was evaluated in 4 of the 14 studies utilizing rating
scales. In brief, these studies reported that teachers found FBA procedures as
accept able and effective in reducing targeted behaviors. One of the studies found
mixed results where half of the staff refused to implement the interventions and the
other half required additional levels of support from the researcher beyond the time
and intensity initially intended by the researcher. Another study used the standardized
Treatment Acceptability Rating Form-Revised (TARF-R) that compared pre- and postintervention perceptions. The TARF-R is a 17-item scale that evaluates perceived reasonableness, effectiveness, intrusiveness, cost, and teacher willingness. Scaled scores range from 17 (i.e., low acceptability) to 119 (i.e., high acceptability). The results of this study reflected a large change (baseline mean=31; posttreatment mean=107) in terms of attitude toward the acceptability and effectiveness of the FBA process.

Effective performance of FBAs requires that the agent is able to understand the underlying behavioral principles in order to define behaviors, collect data, conduct observations, interpret the data to establish a functional relationship, and develop interventions relative to the analysis. Practically speaking, the FBA process must be done in a time and energy efficient manner that does not detract from the other obligations of school staff. In all but one of the cases of this study, the assessment and intervention development was conducted by researchers, which makes assessing practicality difficult and raises concern. Because of the limited involvement of practitioners in these studies, no firm conclusions as to the practicality of FBA were derived from the data.

In summary, this study supported the utility of FBA as a viable method for reducing target behaviors in students with high-incidence disabilities in public-school settings. The analysis did not delve into the various methods of FBA, although there was acknowledgement of the variability in terms of effect and time invested in using the multiple methods of FBA. The results suggested that FBA is acceptable to teachers; however, these conclusions may be suspect given the small sample of studies as well as the relatively low level of practitioner involvement in assessing and
developing interventions. Although the study discussed practicality, there were no substantive data to support or refute claims that FBA is practical for use in the public-school systems. As such, the current meta-analysis, like most of the studies reviewed here, was intended to assist practitioners by establishing guidelines for the effective and efficient use of FBA to develop interventions successful in reducing target behaviors and increasing appropriate replacement behaviors. To expand on the very broad literature base presented here, the current study examined the overall effect of the various FBA methods as well as compared the multiphase and two-level procedures that may guide the FBA process.

Summary

With the richness of diversity and increased homogeneity within the public-school setting comes increased expectations for teacher and student performance that can place additional financial strain on school systems. The number of students who experience socioeconomic pressures from limited familial resources (e.g., access to mental-health and health services) and support affect both academic and social development. Furthermore, financial constraints on families affect accessibility to general and mental-health services, placing a greater burden on educational institutions in meeting these needs. The number of students identified as having learning and emotional and behavioral needs has increased dramatically from 1986 to 2006 (OSEP, 2006). Additionally, the challenges associated with meeting the needs of students with both externalizing (e.g., aggression, self-injurious, and property destruction) and internalizing behaviors (e.g., anxiety and depression) are increasing without the capacity to address the needs (Mayer, 1995; Sugai, Horner, et al., 1999). Furthermore, school personnel are mandated under the reauthorized IDEA (1997) and
IDEIA (2004) to apply the PBS approach and the FBA process within the public-school setting to address pervasive and maladaptive behavior.

To meet these challenges, many districts and schools are implementing PBS, which is a systems approach for providing positive interventions that effect prosocial behavior change (Borgmeier, 2005; Sadler, 2000). Such behavioral changes involve the decrease of problematic behaviors while replacing them with alternative behaviors that will enhance the individual’s life across settings (Carr et al., 1999; Sugai, Horner, et al., 1999). This behavioral orientation differs from punitive methods of behavioral intervention in that remediation is instructional and involves both environmental and interpersonal changes to increase the occurrence of appropriate behavior. These multicomponent interventions generally are derived from a form of FBA and address antecedent and setting event manipulations, consequence strategies, instruction of alternative skills, and global lifestyle interventions (Carr et al., 1999; Hanley et al., 2003). Because the environment influences behavior, the success of any multicomponent intervention is dependent upon the ability to manipulate the context of the behavior or behaviors.

Functional behavioral assessment is a systematic method of operationally defining pervasive maladaptive behaviors and the predictable antecedents and consequences that maintain the behavior over a given period of time (Gresham, Quinn, & Restori, 2001; Iwata, Dorsey, Slifer, Baumann, & Richman, 1994). In the educational setting, the assessment may consist of one or a combination of FBA methods, which include indirect, direct descriptive observations, and functional assessment.
The first method is referred to as the indirect assessment, which may consist of interviews, rating scales, review of archival records, review of medical history and treatment, or the functional assessment interview (FAI). The second method, the direct descriptive observation, may involve an antecedent and behavior (AB) or antecedent, behavior and consequence (ABC); frequency; temporality; intensity; or permanent byproduct record. The final FBA method is FA, which can be either an experimental manipulation or brief FA.

The data from these assessments are then analyzed and translated into a testable hypothesis, which must be validated through further assessment across settings (O’Neill et al., 1997). Provided the hypothesis yields reliable occurrence and nonoccurrence of the target behavior given the prescribed conditions and contexts, a behavior intervention plan may be developed. The plan is intended to be multidimensional in design in that the focus becomes the following: teaching alternative appropriate behaviors, modifying the learning environment to enhance the opportunities to demonstrate the alternative behaviors, reinforcing the occurrence of replacement behavior, reducing direct and indirect variables that encourage the reoccurrence of the target behaviors, and developing strategies to respond to the occurrence of target behaviors (Carr et al., 1999; Horner, 2004).

The description above, however, is a model that is well established, yet highly debated in the literature and research community because of the nuances inherent in any assessment process (e.g., methods of analyzing behavioral events stimulus-response, reinforcement-based interventions, and antecedent-behavior-consequence methods; Hanley et al., 2003). Some researchers have suggested that more extensive FBAs produce more effective intervention plans (Quinn, 2000). Other members of the
research community, however, have supported the use of indirect and direct methods as reasonable means to calculate appropriate function-based interventions plans (Doggett, Edwards, Moore, Tingstrom, & Wilczynski, 2001; O’Neill et al., 1997). In light of the apparent controversy surrounding general guidelines for conducting an FBA and legal mandates associated with this practice, there is a clear need to analyze systematically the context for these practices, the various methods and the possible relationship with behavioral interventions. Furthermore, the research-to-practice gap has researchers expanding the traditional criteria for evaluating an FBA to include social validity and lifestyle change indicators in an effort to increase the application of FBA and PBS with more typical agents in more typical settings (Carr et al., 1999; Kincaid et al., 2002).

Therefore, the purpose of this study was to continue the research reviewed in this chapter by conducting a meta-analysis of single-case studies that examined behavioral change between the baseline and intervention phases. As with the Reid and Nelson study (2002), this analysis investigated the average effect of FBA-based interventions for students with high-incidence disabilities; however, unlike many of the studies in this literature review, the current research sought to identify and compare the effects of the various methods of FBA. Furthermore, the analysis of the model used in conducting the assessment (i.e., two-level or four-phase FBA; see Figures 2 & 3 in Chapter 1) and the order of methods used in completing the FBA may provide insights that can address many of the acceptability, practicality, and utility issues (Carr et al., 1999; Kincaid et al., 2002; Reid & Nelson, 2002).
CHAPTER III

METHODOLOGY

The purpose of this study was to analyze the effect sizes for the various methods of conducting a functional behavioral assessment (FBA) and the resulting behavior intervention plan (BIP) intended to address pervasive and maladaptive behaviors within the public-school setting. Further examination of the relationship between the FBA process and BIP outcomes may provide practitioners with guidelines for more effective and efficient procedures for assessment and the implementation of interventions. This chapter contains details regarding the research design, strategies for literature search, inclusion criteria, variables, effect-size measure, instrumentation, and coding procedures.

Method

Given the large number of studies dedicated to both FBA as a process (i.e., Ervin et al., 2001; Hanley et al., 2003; Reid & Nelson, 2001) and behavior intervention plan (BIP) as an outcome of this process (Durand & Carr, 1985; Hanley et al., 2003), this study analyzed only pervasive and maladaptive behaviors among school-age students within the public-school setting. A meta-analysis may be the most straightforward manner to measure the relationship between the two in quantifiable terms and synthesize the results into a more practitioner-friendly model. The meta-analysis procedure allowed the researcher to synthesize the results of many studies in terms of the overall effect of the treatment. In this study, the independent variable was the FBA method and the disability type and the dependent variable was the effect size of the behavioral difference. The results of this statistical process were compared by the variables in the methods used in conducting the FBA.
For the purpose of this study, FBA methods were divided into three primary categories: indirect, direct descriptive observation, and functional assessment. Many studies reflect a multiple methods approach in conducting a FBA (Dahlstrom, 2003). There are seven possible combinations or individual methods that may appear within a study: Indirect only, Direct descriptive observation only, Functional Assessment (FA) only, Indirect + Direct descriptive, Indirect + FA, Direct descriptive + FA, and Indirect + Direct descriptive + FA.

These methods, or combinations of methods, when distilled further, may reveal practical techniques useful in the development of a FBA and a BIP. Indirect methods of assessing include archival reviews, ratings scales, and functional assessment interview (FAI; Horner & Carr, 1997). The Direct descriptive observation method may involve an antecedent-behavior (AB) or antecedent-behavior-consequence (ABC), frequency, temporality, intensity, or permanent byproducts technique for collecting data. The most time intensive method is the functional assessment, which can be a functional analysis with experimental manipulation (Iwata et al., 1994), a brief functional analysis (Northrup et al., 1991), or some of other preexperimental data analysis of the hypothesis and intervention.

To assist practitioners in conducting a FBA and implementing a BIP, this study used of meta-analytic procedures to investigate the methods of assessment used in the FBA process. The current research examined the effect sizes of FBA-based interventions calculated using pretest and posttest results in terms of behavioral change. The various sets and subsets of indirect, direct descriptive, and FA methods of FBA as well as disability type were compared through 2 one-way analysis of
variance (ANOVA). A notable difference between the various methods may provide practitioners guidance in the FBA process.

Strategies for Literature Search

A comprehensive search of published and unpublished material using bibliographic databases was used to locate literature for consideration in this meta-analysis, which included PsychINFO, Dissertation Abstracts International, Dissertation Abstracts Online, PsycLit, Ed Source, Educational Resources Information Center (ERIC), U.S. Government Printing Office publications, National Criminal Justice Reference Services, and the Office of Special Education Programs. Additionally, a review was conducted of textbooks, related journals, previous meta-analyses, bibliographies of other studies, and other literature reviews. Professional organization publications that are not peer-reviewed, organization reports and papers, and conference papers were screened for additional references. Members of Positive Interventions Network of Trainers (PENT) of California, Applied Behavior Analysts (ABA), and California Applied Behavior Analysts (Cal-ABA) were contacted through a list serve to identify other related articles or material that may or may not be published. Furthermore, references included in articles identified as being relevant to the topic were reviewed as additional potential sources.

Database searches and other related queries included the following keywords: behavioral analysis, functional analysis, functional assessment, Functional Analysis Assessment, Functional Behavioral Assessment, at-risk, disruptive behavior, aggressive behavior, bullying, fighting, tantrum, school violence, externalizing behaviors, self-injurious behavior, terrorist threat, inattention, acting out, IDEA discipline, IDEIA discipline, reinforcement, Positive-behavior supports, applied
behavioral analysis (ABA), classroom management, behavioral, behavioral interventions, autism, autistic spectrum disorder, emotional disturbance, conduct disorder, behavioral disorder, oppositional -defiant disorder, school-wide discipline, assertive discipline, social skills, citizenship, academic interventions, Attention Deficit Disorder with Hyperactivity (ADHD), Attention Deficit Disorder (ADD), mental retardation, and Down’s Syndrome.

The studies considered for inclusion in this study dated back from 1975 to the present. The P.L. 92-142 or the Education of All Handicapped Children Act (EAHC) was enacted in 1975, which is considered the period in which the application of functional assessments and positive-behavioral-intervention strategies began to emerge within the public-school setting. The EAHC also introduced the concept of least restrictive environment. This range of dates allowed for easier comparison among interventions that were developed without utilizing FBA as the primary method for designing the intervention. Articles published prior to 1975 were considered for background information but were not included directly in the analysis.

Inclusion Criteria

Initial consideration for inclusion in the study was similar to the criteria proposed by Wilson, Lipsey, and Derzon (2003) for their comparison study, which were as follows:

1. Study reported in English, published in 1975 or after, and involved school-age students at least partly assessed in a public school.
2. Study assessed at least one intervention effect for at least one outcome variable that represented pervasive and maladaptive behavior.
3. Study used either of the following designs:
a. Single-case study design in which measures of at least one qualifying outcome variable was taken before and after intervention on the same subjects, including one-group designs and multiple-group designs involving different interventions.

b. These studies included one or more of the FBA methods detailed in the current study: indirect, direct descriptive, or functional assessment.

The criteria were broad to allow for a greater inclusion of studies in the meta-analysis that could then be further disaggregated into disability, setting, model, method of FBA, intervention, behavior, individual assessing, individual implementing, and other related supportive factors. This approach to inclusion may help to increase the validity of the search by minimizing retrieval bias and by allowing for greater representation among the sample of students across the usual demographic characteristics such as gender, ethnicity, language, age, socioeconomic status, and the like.

Unlike many of the sample studies, this meta-analysis excluded subjects with less frequently occurring disability categories of autistic spectrum disorder, mental retardation, multiple disabilities, and other legally defined low-incidence disabilities defined by funding statutes as visually impairment, deaf-blind, hearing impairment, orthopedic impairment, and traumatic brain injury. Instead, this study focused on students with high-incidence disabilities who constitute 83% of all students served through special education in the state of California (California Department of Education, 2005). These high-incidence disabilities include specific learning disabilities, emotional disturbance, other health impairment, and speech and language disorders (Broussard & Northrup, 1995). The study included cases with students
identified as having high-incidence disabilities with Section 504 plans as the Civil Rights protections afforded under this statute parallel the IDEA (1997) with regard to discipline.

Variables

The dependent variable for this study was the difference, if any, between the baseline- and intervention-phase means divided by the standard deviation of the baseline measure, which was the corresponding effect size.

The independent variables were the type of method of functional behavioral assessment used to evaluate target behavior, the behavior intervention plan that resulted from the various methods of FBA, and the type of disability.

Other relevant variables examined in this study included disability, FBA method technique, sequential information related to the use of multiple methods and techniques, as well as supportive factors that may contribute to the development and implementation of the assessment and intervention such as school-wide supports, assessor, implementer, topography of the behavior, and training.

Effect-size Measure

The purpose of this meta-analysis was to assess the overall effect of interventions intended to address pervasive and maladaptive behaviors and to analyze the process practitioners and researchers use in developing these interventions. To be included in this meta-analysis, the single-case studies must comprise at least one baseline and intervention phase measure and provide the standard deviation of the baseline in order to calculate the effect size. This method of analysis allowed for the inclusion of studies that used the various types of single-case study designs such as reversal, simple single-case, alternating treatment, and multiple-baseline. One effect size was
calculated for each relevant dependent variable per individual in the analysis of the single-case studies. The effect-size index for Cohen’s (1988) \( d \) was interpreted as follows: .2 (small), .5 (medium), and .8 (large).

The effect-size statistic for this study was defined as the difference between the intervention-phase mean and the baseline-phase mean change divided by the standard deviation of the baseline phase (Busk & Serlin, 1992; Glass, 1978) or

\[
d_i = \frac{\bar{Y}_i - \bar{X}_i}{s_i}
\]

where \( \bar{X}_i \) was the baseline-phase mean and \( \bar{Y}_i \) was the intervention-phase mean, and \( s_i \) was the standard deviation of the baseline phase across for Subject (i).

Coding Sheet

The researcher and a practitioner familiar with applied behavioral analysis (ABA), positive behavior support (PBS), FBA, BIP, special education, and school psychology independently coded the studies. The practitioner was trained in the identification of the variables, the coding process, and documentation through a commonly-used spreadsheet application. Included in the training was a practice coding trial of 25 studies conducted in a clinical setting as opposed to public-school setting. The researcher and practitioner coded 10 studies related to the pharmacological treatment of attention deficit disorder hyperactive-type (ADHD) to rehearse the coding of environmental features of a public-school setting such as elements of the PBS system of support or staff implementing. Due to the limited number of studies that meet the inclusion criteria, actual studies that meet the criteria were not used for practice and training purposes. This practice coding trial was used to compare intercoder reliability and to evaluate the variables to be coded prior to
conducting the actual metanalysis. Any variables perceived to be ambiguous were revised.

Effective reliability of the intercoder correlation for both the training and the actual study was calculated using the Spearman-Brown product moment correlation coefficient, where \( R \) was the effective reliability result, \( n \) was the number of judges, and \( r \) was the mean reliability calculated based upon point-by-point agreement among all \( n \) judges (i.e., mean of \( n(n-1)/2 \); Rosenthal, 1991). The Spearman-Brown formula appears as follows:

\[
R = \frac{nr}{1 + (n - 1)r}.
\]

The studies were coded by each rater according to the following characteristics and rationale (see Appendix B):

1. **Author of the study**
2. **Date of study**
3. **Single-case study design** is necessary in determining the effect size, validity, and reliability for each study. In single-case studies, the baseline condition is A, the intervention condition is B, and other intervention conditions follow ascending alphabetical order (e.g., C, D, etc.).
   a. Reversal (e.g., A-B-A)
   b. Simple single-case (e.g., A-B)
   c. Alternating treatment (e.g., A-B-C-A)
   d. Multiple-baseline (e.g., \( A_a(\text{anxiety})-B_m(\text{medication}) \) and \( A_d(\text{depression})-B_c(\text{Group counseling}) \).
4. Source type is helpful in comparing primary sources for related articles included in the study. The sources d-f were included in this analysis and screened for additional sources.
   a. Peer-reviewed journal
   b. Dissertation
   c. Textbook
   d. Unpublished work
   e. Professional journal (not peer-reviewed)
   f. Published or unpublished case study

5. Gender of the individual receiving the treatment
   a. Male
   b. Female
   c. Not specified

6. Socioeconomic status of subject
   a. Low
   b. Middle
   c. High
   d. Not specified

7. Geographic location of study

8. Age of group, years

9. Age of group, range

10. Ethnicity of the individual receiving the treatment
    a. European American
    b. Hispanic American
c. African American
d. Asian American
e. Native American
f. Other
g. Not Specified

11. *Intervention type* is the resultant of the various FBA methods.
   
a. Counseling
b. Environmental manipulation
c. Tangible reinforcement
d. Social reinforcement or socially-based
e. Punishment
f. Consequence only-no positive reinforcement
g. Removal of individual
h. Removal of undesired item or task
i. Token economy
j. Adult praise or attention
k. Other (Extinction, DRO, etc.)

12. *Behavior* must be operationally defined. The behavior, behaviors, or class of behaviors are target of the FBA process and addressed through the BIP.
   
a. Physical aggression
b. Verbal aggression
c. Self-injurious
d. Property destruction
e. Possession of dangerous weapon  
f. Possession of controlled substance  
g. Inattention  
h. Sensory seeking or avoiding  
i. Other operationally defined behavior

13. *Function of behavior* is a pivotal factor in determining the appropriate intervention. The predictive value of a FBA method is often evaluated using the appropriate identification of the function.

   a. Escape, protest, or avoid  
   b. Get or obtain

14. *Personnel conducting assessment* is important to this study as one of the primary reasons for a research-to-practice gap is that the majority of studies in this field are conducted by someone other than staff that have daily contact with the student.

   a. Classroom practitioner (teacher, classroom staff, etc.)  
   b. School psychologist  
   c. School behaviorist  
   d. Other school specialist  
   e. Researcher  
   f. Research/university affiliate, for example, research assistant, assisting researcher, graduate student, etc.

15. *Individual supervising implementation of assessment results* may be a relevant factor that influences the effectiveness of the intervention and the social validity of the FBA process.
a. Typical agent
   i. Classroom practitioner (teacher, classroom staff, etc.)
   ii. School psychologist
   iii. School behaviorist
   iv. Other school specialist
   v. Parent

b. Atypical agent
   i. Researcher
   ii. Research/university affiliate

16. Personnel implementing intervention may differ from the individual conducting the FBA and supervising the process.

   a. Typical agent
      i. Classroom practitioner (teacher, classroom staff, etc.)
      ii. School psychologist
      iii. School behaviorist
      iv. Other school specialist
      v. Parent

   b. Atypical agent
      i. Researcher
      ii. Research/university affiliate

17. Rated fidelity in implementation

   a. Very poor
   b. Poor
   c. Average
d. Good

e. Very good

f. Not reported by researcher

18. Level of post intervention follow-up to evaluate fidelity of implementation

a. No follow-up reported

b. 0-3 months

c. 3-6 months

d. More than 6 months

19. Assessment model describes the overall framework used in conducting the FBA and potentially implementing the BIP.

a. Multiphase model

b. Two-level

c. Other model

d. None evidenced

20. Primary disability

a. Specific learning disability

b. Speech and language impairment

c. Orthopedically impaired

d. Other health impaired

e. Autism

f. Mental retardation

g. Emotional disturbance

h. Visually impaired

i. Blind
j. Deaf/Hearing impairment
k. Visually and hearing impaired
l. Multiple-disabilities
m. Section 504 eligibility only

21. Method of FBA
  a. Direct (D) only
  b. Indirect (I) only
  c. Functional assessment (FA) only
  d. D+I
  e. I+D
  f. D+FA
  g. FA+D
  h. I+FA
  i. FA+I
  j. I+D+FA
  k. D+I+FA
  l. FA+D+I
  m. FA+I+D
  n. I+FA+D
  o. D+FA+I

22. Baseline phase mean

23. Intervention phase mean

24. Standard deviation of baseline phase
The researcher and the practitioner trained in the meta-analytic procedures coded the included studies based upon a variety of variables detailed in Appendix B. Two practice trials were conducted using studies that were not included in the meta-analysis, which resulted in 91% and then 98% agreement using the Spearman-Brown product moment correlation coefficient (Rosenthal, 1991). With an acceptable level of effective reliability in the practice trials, the researcher and practitioner coded the data sets of the included studies calculating the means and standard deviations. The effect sizes were calculated by subtracting the baseline-phase mean from the intervention-phase mean divided by the standard deviation of the baseline phase (Busk & Serlin, 1992; Glass, 1978). The effective reliability for the meta-analysis was 98% on the initial coding and analysis. When there was disagreement between the raters, the researcher and practitioner reviewed and recalculated the statistical information to enhance accuracy of reporting the results. Discrepancies in the data were attributed to the visual inspection of the data points presented in the single-case studies and in one case a missed key stroke while inputting the data.

Data Analysis

Specifically, this study examined the following research questions:

1. What was the average effect size for each of the various FBA-based methods and interventions for students with high-incidence disabilities, such as learning disabilities, attention deficit disorder with hyperactivity, or emotional or behavioral disorders?

2. Which FBA methods yielded the greatest average effect sizes and how did the
various methods compare with each other across the various high-incidence
disabilities?

3. If multiple FBA methods were used, was there a difference in the average
effect sizes based on the order in which the methods were used?

4. Did the assessment and intervention process follow a specific framework,
such as a multiple-phase or two-level model, from beginning to end? If so, was there
a difference in average effect size for the interventions developed within a given
framework?

The first research question related to the effectiveness of FBA-based
interventions and possible average effects of the intervention was evaluated by
calculating the standardized difference in means for the baseline and intervention
phases divided by the baseline standard deviation. Homogeneity of variance of effect
sizes was assessed using a Hedges’ (1981) $Q$ test to determine whether the observed
difference in effect sizes is due to sampling error. If not homogeneous, attempts were
made by eliminating extreme effect sizes to obtain a homogenous set of studies and
the overall effect size. Assuming the calculated Q statistic for the individual effect
sizes exceeds the critical value for the upper limits of chi-squared distribution, the
individual effect sizes can be pooled and average effect sizes reported. The
confidence interval was calculated so that statistically significant differences from
zero could be assessed. The effect size for the dependent variable or variables was
calculated and compared using Cohen’s $d$ index (1988). A larger effect size would
suggest a greater or more positive effect in terms of reducing the intensity and
duration of the target behavior, whereas, a smaller effect would reflect less of a
decrease in the target behavior.
The second research question involved comparing and evaluating the independent variables, the various FBA methods and disability types, and the dependent variable, the effect size of any behavioral difference. The effect sizes were planned to be analyzed using a two-way analysis of variance (ANOVA) for differences across the various methods and between the different types of high-incidence disabilities. These multiple methods included the following: Indirect only, Direct descriptive observation only, Functional Assessment (FA) only, Indirect + Direct descriptive, Indirect + FA, Direct descriptive + FA, and Indirect + Direct descriptive + FA. This analysis may assist practitioners in deciding which method may provide the most effective intervention given a specific population. Future research might evaluate the effectiveness of the different methods given a certain behavioral topography or multifunction behaviors. Because of missing cell information, 2 one-way ANOVAs were performed.

Where the second research question examined the effectiveness of one method to another, the third question explored whether there were any implications for a given sequence or order of assessment methods used by a practitioner. In other words, assuming there was some difference between the methods, was there any significant difference in the effectiveness of a BIP based upon an FBA that varies the order of the assessment method for example from the standard Indirect + Direct descriptive + FA to Direct descriptive + Indirect + FA. To put this question into more practical terms, the practitioner might want to know if it is better to interview the teacher and parents first and then conduct the observations or is it better to do the observation first. This analysis compared the resulting effect sizes for difference of means relative
to the order of the assessment method. The comparison of order may assist practitioners in conducting more timely and accurate assessments.

The fourth question addressed two established frameworks for conducting a FBA. The first method is sequential in following the four phases, descriptive, interpretive, hypothesis testing, and intervention and implementation (Ervin et al., 2001). The assessor determines the methods of assessment and data collection during the evaluative process as information is refined narrowing the hypothesis test and the intervention to increase precision. This method may require a greater initial investment of time in conducting the assessment; however, the results are effective in most cases (Ervin et al., 2001). In contrast, the second model developed by Sugai, Lewis-Palmer, and Hagan-Burke (1999) provides greater flexibility in terms of the intensity and procedures used. This model, often referred to as the two-leveled approach, matches the intensity of the assessment, data collection, and intervention with the level of the behavior. In this dynamic manner, the assessor can test assumptions and environmental manipulations in response to behavioral changes of the student. Independent-samples t tests were used to compare the two models. This analysis may be a key in establishing guidelines that consider not only the importance of an effective evaluation from a clinical standpoint but also the social acceptability of the FBA process within the school system.
CHAPTER IV

RESULTS

The purpose of this study was to conduct a meta-analysis of single-case studies that examined pervasive and maladaptive behaviors demonstrated by students with high-incidence disabilities within the public-school setting. The results of the meta-analysis detailed in this chapter are derived from a literature search that included a review of 357 titles and abstracts from technical reports as well as published and nonpublished works related to functional behavioral assessment and behavioral interventions written between 1975 and 2006. Of the reviewed studies, 91% involved students with low-incidence disabilities such as mental retardation, autistic spectrum disorder, pervasive developmental disorders, and combinations thereof or involved preschool-aged students and, therefore, were excluded from this current study. Although 9% of the studies included high-incidence disabilities such as specific learning disabilities (SLD), emotional disturbance (ED), conduct disorders (CD), and so forth, only 4% or 15 studies were conducted at least in part within the public-school setting and met the inclusion criteria for this study (see Chapter 3). Five of the included 15 studies were nonpublished technical papers, dissertations, or theses, and the remainder of the studies were from peer-reviewed journals. The other 5% of the studies with high-incidence disabilities were excluded from this meta-analysis because the research was conducted in clinical settings. Studies that met the inclusion criteria had publication dates between 1992 and 2006, which suggests a more recent trend toward analyzing pervasive and maladaptive behavior with public-school students having high-incidence disabilities.
Depicted in Table 4 are the characteristics of the 15 studies that included 28 individuals who demonstrated pervasive and maladaptive behaviors. These samples were selected based upon the operationally-defined behaviors, disability type, and the setting in which the assessment and intervention occurred.

Table 4

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>( f )</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Source</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Journal of Applied Behavior Analysis</td>
<td>4</td>
<td>27</td>
</tr>
<tr>
<td>School Psychology Review</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>Journal of Behavior Interventions</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>Education and Treatment of Children</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>National Association of School Psychologists Review</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>Dissertation</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>Masters Thesis</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>Technical Paper</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>23</td>
<td>82</td>
</tr>
<tr>
<td>Female</td>
<td>5</td>
<td>18</td>
</tr>
<tr>
<td><strong>Disability type</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotionally Disturbed (ED)</td>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>ED and ADHD</td>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>Specific Learning Disability (SLD)</td>
<td>5</td>
<td>18</td>
</tr>
<tr>
<td>SLD and ED</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>SLD and ADHD</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>OHI (ADHD)</td>
<td>5</td>
<td>18</td>
</tr>
<tr>
<td>CD</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>No identified disability</td>
<td>5</td>
<td>18</td>
</tr>
<tr>
<td><strong>Age Range</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-7 years</td>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>8-10 years</td>
<td>12</td>
<td>43</td>
</tr>
<tr>
<td>11-13 years</td>
<td>12</td>
<td>43</td>
</tr>
<tr>
<td><strong>Agent</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atypical (e.g., researcher, graduate assistant, etc.)</td>
<td>8</td>
<td>28</td>
</tr>
<tr>
<td>Typical (e.g., teacher, instructional assistant, etc.)</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Both Typical and Atypical</td>
<td>18</td>
<td>64</td>
</tr>
</tbody>
</table>
The students ranged between 5 to 13 years of age. Student disability types included SLD, ED, CD, ADHD, and OHI with many students having comorbid diagnoses (e.g., ED and ADHD). The most prevalent disability type categories were SLD, ADHD, and those who exhibited pervasive and maladaptive behaviors who were not identified as being eligible for special-education services due to a disability (18% each) followed by ED and ED with ADHD (14% each).

The individual with a diagnosed condition of CD was identified uniquely as having needs that necessitated a Section 504 Accommodation Plan under the ADA. Nearly 86% of the individuals were between 8 and 13 years of age. Of those individuals included in this study, 82% were males.

A review of the agents or individuals conducting assessment and implementing the intervention found that 64% of the FBAs involved atypical and typical agents. Atypical agents are defined as researchers, clinical behavioral specialists, research assistants, and so forth who most typically conduct FBAs and implement the intervention plans. Typical agents are often times teachers, parents, instructional assistants or aides, and students, if this meta-analysis is an indication of individual-centered assessments and interventions. Several studies involved the individual student through indirect methods of evaluation, such as reinforcement inventories, interviews, and, in a few cases, a review of antecedent and consequence conditions after the individual demonstrated the target behavior. The practitioner referred to this ABC self-reflection as a “behavioral autopsy” because the data collection forms were used in an indirect manner after the behavior had occurred as opposed to in the customary direct observation form. Another study capitalized on
social reinforcement by employing peers to assist with a self-monitoring plan and contingent reinforcement.

The results of the literature review and coding were used to examine the four research questions posited in this meta-analysis that explored the effects of various FBA methods used in developing interventions to address pervasive and maladaptive behaviors in students with high-incidence disabilities such as ADHD, SLD, ED, CD, and combinations of the aforementioned conditions. Furthermore, this study compared the relationship between the various FBA methods, the order of conducting the assessment, and the framework or model followed in the evaluation of the behavior and intervention in addressing the following research questions:

1. What is the average effect size for each of the various FBA-based methods and interventions for students with high-incidence disabilities, such as learning disabilities, attention deficit disorder with hyperactivity, or emotional or behavioral disorders?

2. Which FBA methods yielded the greatest average effect sizes and how did the various methods compare with each other across the various high-incidence disabilities?

3. If multiple FBA methods were used, is there a difference in the effect sizes based on the order in which the methods were used?

4. Did the assessment and intervention process follow a specific framework, such as a multiple phase or two-level model, from beginning to end? If so, was there a difference in effect sizes for the interventions developed within a given framework?
The effect-size statistic was calculated as the difference between the intervention-phase mean and the baseline-phase mean change divided by the standard deviation of the baseline phase (Busk & Serlin, 1992; Glass, 1978) or

\[ d_i = \frac{\bar{Y}_i - \bar{X}_i}{s_i} \]

where \( \bar{X}_i \) was the baseline-phase mean and \( \bar{Y}_i \) was the intervention-phase mean, and \( s_i \) was the standard deviation of the baseline phase across for Subject (i).

For this study, effect sizes are calculated for the individual in the single-case studies and not the distinct studies are compared using Cohen’s \( d \) as seen in Table 5. Effect sizes were calculated for individuals instead for the studies to allow for further analysis of some of the unique individual characteristics, such as disability type or mixed methods of assessment, which might not be captured through examination of the single-case study. Assuming normal distribution, percent of nonoverlap, refers to the degree to which there is no overlap of the phenomenon present between groups being compared (Cohen, 1988). The effect-size statistic represents a fraction of a SD from the mean (e.g., a .5 ES= ½ of a SD; 2.0= 2 SDs). Given Cohen’s \( d \) standard, it is possible to extrapolate that an effect size of 3.0 has percentile standing of nearly 99% with a percent of nonoverlap of nearly 90%, which is an extremely large effect size. Because the overall range of the effect sizes is from small (0.14) to questionably large (7.68) and an overall mean of 3.09, a brief review of Cohen’s \( d \) may be not only helpful but also necessary in interpreting the magnitude of the results under each research question.
Table 5

Summary of Cohen’s $d$ Index

<table>
<thead>
<tr>
<th>Cohen's $d$ Index</th>
<th>Effect Size</th>
<th>Percentile Standing</th>
<th>Percent of Nonoverlap</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.0</td>
<td>98</td>
<td>81.1</td>
<td></td>
</tr>
<tr>
<td>1.9</td>
<td>97</td>
<td>79.4</td>
<td></td>
</tr>
<tr>
<td>1.8</td>
<td>96</td>
<td>77.4</td>
<td></td>
</tr>
<tr>
<td>1.7</td>
<td>96</td>
<td>75.4</td>
<td></td>
</tr>
<tr>
<td>1.6</td>
<td>95</td>
<td>73.1</td>
<td></td>
</tr>
<tr>
<td>HUGE</td>
<td>1.5</td>
<td>93</td>
<td>70.7</td>
</tr>
<tr>
<td></td>
<td>1.4</td>
<td>92</td>
<td>68.1</td>
</tr>
<tr>
<td></td>
<td>1.3</td>
<td>90</td>
<td>65.3</td>
</tr>
<tr>
<td></td>
<td>1.2</td>
<td>88</td>
<td>62.2</td>
</tr>
<tr>
<td>VERY LARGE</td>
<td>1.1</td>
<td>86</td>
<td>58.9</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>84</td>
<td>55.4</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>82</td>
<td>51.6</td>
</tr>
<tr>
<td>LARGE</td>
<td>0.8</td>
<td>79</td>
<td>47.4</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>76</td>
<td>43.0</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>73</td>
<td>38.2</td>
</tr>
<tr>
<td>MEDIUM</td>
<td>0.5</td>
<td>69</td>
<td>33.0</td>
</tr>
<tr>
<td></td>
<td>0.4</td>
<td>66</td>
<td>27.4</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>62</td>
<td>21.3</td>
</tr>
<tr>
<td>SMALL</td>
<td>0.2</td>
<td>58</td>
<td>14.7</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
<td>54</td>
<td>7.7</td>
</tr>
<tr>
<td></td>
<td>0.0</td>
<td>50</td>
<td>0.0</td>
</tr>
</tbody>
</table>

A $Q$ statistic testing for homogeneity of effect sizes was calculated first for all of the individual effect size, which allows for the report of average effects based on the pooled individual effect sizes for the various FBA methods, disability types, and so forth. The $Q$ statistic for the 28 individuals, estimating the homogeneity of effect-size estimates was not statistically significant ($Q = 22.3; df = 27$) suggesting homogeneity of these effect-size estimates. Similarly, $Q$ statistics were calculated on the effect sizes grouped for assessment method ($Q = 20.3; df = 27$), disability type ($Q = 9.12; df = 7$), and FA model ($Q = 1.61; df = 1$), length ($Q = 4.98; df = 2$), and type
(Q = 2.7; df=2). All of the subsequent Q statistics were found to be nonsignificant, and, therefore, homogeneity was assumed.

**Research Question 1**

The frequency and effect sizes and standard deviations for the FBA methods in the 15 studies conducted with 28 individuals with high-incidence disabilities are summarized in Table 6. The effect size means ranged from a low of 2.34 to a high of 4.07 with a grand mean of 3.17 and overall SD of 2.14. Preliminary analysis of the usage of the various FBA methods and the average effect sizes of the corresponding methods suggested that there were statistically and practically significant effects for FBA-based methods and interventions.

The more comprehensive and presumably more-time-consuming Indirect+Direct+FA method was most prevalent in the studies and was second highest (ES=3.14) in terms of average effect size. The greatest average effect size (ES=4.07) resulted from two studies that utilized Direct+Indirect methods that did not make use of FA. These results should be viewed, however, with caution given the small number of cases (f=2) and that the data were retrieved from a master’s level thesis and a technical paper written by the practitioner. Comparatively, Direct+FA had the lowest average effect size (ES=2.34), which was still a considerable effect size given the standard for Cohen’s $d$ (i.e., .2 = small effect, .5 = medium effect, and .8 = large effect). Neither the order in which the method of assessment was conducted nor the disability type are considered in Table 6. These results suggest that all of the assessment methods analyzed in the literature about addressing pervasive and maladaptive behaviors of students with high-incidence disabilities within public-school settings yielded highly effective levels of behavioral change over baseline.
Table 6

Average Effect Sizes for the Various FBA-based Methods and Interventions

<table>
<thead>
<tr>
<th>Assessment Method</th>
<th>f</th>
<th>Average ES</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Functional assessment (FA) only</td>
<td>4</td>
<td>3.13</td>
<td>2.34</td>
</tr>
<tr>
<td>Direct+Indirect</td>
<td>2</td>
<td>4.07</td>
<td>2.02</td>
</tr>
<tr>
<td>Direct+FA</td>
<td>5</td>
<td>2.34</td>
<td>2.27</td>
</tr>
<tr>
<td>Indirect+Direct+FA</td>
<td>17</td>
<td>3.14</td>
<td>1.90</td>
</tr>
<tr>
<td>Total</td>
<td>28</td>
<td>3.17</td>
<td>2.14</td>
</tr>
</tbody>
</table>

Although all of the methods demonstrated a positive effect in terms of behavioral change, the first research question involved the comparison of average effect sizes among the FBA methods found in the literature with the high-incidence student population. Of the seven FBA methods utilized in the greater body of research across all disability types, the FBA methods encountered in the literature search for high-incidence disabilities included FA only, Indirect+Direct, Direct+FA, and Indirect+Direct+FA. Twenty-six of the 28 individuals received at least some form of FA as part of the assessment. Unlike studies involving students with low-incidence disabilities, the Indirect only, Direct descriptive only, and Indirect + FA methods were noticeably absent from the FBA methods for students with high-incidence disabilities. The absence of Indirect and Direct only methods in the literature may be related to the comparatively low number of atypical agents conducting and implementing the FBA and interventions who may be more inclined toward these less quasi- and experimental methods.

The data from the single-case studies reviewed in this meta-analysis appeared to have been collected in manner to ensure that the results were independent of each other, and, therefore, the assumption of independence was not violated. Prior to the main data analysis, the data were examined through visual inspection for outliers and
normality of distribution. Tests of skewness, kurtosis, and normality indicated that there was no statistically or practically significant departure from normality as all of the tests resulted in an absolute value of less than one. Tests of homogeneity of the average effect sizes revealed relatively stable variability across the studies, thereby, meeting the homogeneity of variance assumption.

**Research Question 2**

The second research questions compared the various FBA methods across the various high-incidence disabilities, such as SLD, ED, CD, and so forth. As can be seen in Table 7, the disaggregated effect sizes for assessment method and disability type ranged from 0.87 to 6.60 with a grand mean ES of 3.10. Because of the relatively few individuals identified in the literature as demonstrating pervasive and maladaptive behaviors and having high-incidence disabilities or no disability at all, there are no data a number of cells necessary for the analysis.

**Table 7**

Average Effect Sizes for the Various FBA-based Methods and Interventions Across Disability Types

<table>
<thead>
<tr>
<th>Disability</th>
<th>Functional assessment</th>
<th>Direct+ Indirect</th>
<th>Direct+FA</th>
<th>Indirect + Direct +FA</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ES</td>
<td>f</td>
<td>ES</td>
<td>f</td>
<td>ES</td>
</tr>
<tr>
<td>ED</td>
<td>0.87</td>
<td>1</td>
<td>3.14</td>
<td>3</td>
<td>2.01</td>
</tr>
<tr>
<td>ED and ADHD</td>
<td>1.32</td>
<td>1</td>
<td>6.60</td>
<td>3</td>
<td>3.96</td>
</tr>
<tr>
<td>SLD</td>
<td>3.28</td>
<td>5</td>
<td>3.28</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>SLD and ED</td>
<td>3.75</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLD and ADHD</td>
<td>5.50</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OHI (ADHD)</td>
<td>1.29</td>
<td>1</td>
<td>2.38</td>
<td>3</td>
<td>1.68</td>
</tr>
<tr>
<td>CD</td>
<td>2.01</td>
<td>1</td>
<td>2.01</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>None identified</td>
<td>2.64</td>
<td>1</td>
<td>5.70</td>
<td>1</td>
<td>1.70</td>
</tr>
<tr>
<td>Total</td>
<td>2.52</td>
<td>4</td>
<td>4.07</td>
<td>2</td>
<td>2.27</td>
</tr>
</tbody>
</table>
Originally, the design of this meta-analysis called for a two-way analysis of variance (ANOVA) with assessment method and disability as the main effects to examine interaction between the variables. Due to the absence of data that resulted in the empty cells a 4X8 two-way ANOVA could not be conducted. Table 7, however, does present practically significant results and patterns for further analysis across disability types.

A cursory analysis of these data indicated that the more comprehensive Indirect+Direct+FA method was employed most frequently and was the most effective method of developing intervention across all disability types ($f=18$; $ES=3.21$). As noted in the previous research question, the Direct+Indirect method did provide the largest marginal mean for method ($ES=4.07$); however, these results should be viewed with caution given the small number of cases and that the cases reflect two different disability types, further weakening any generalization of the results. The other two methods FA and Direct+FA shared similar numbers and nearly were comparable as far as levels of practical significance.

With regard to specific disability types, all of the students with ED, often perceived as the most challenging diagnosis to address within the educational system, or a combination of ED with another disability ($f=16$), all received a method that included FA with Direct or with Direct+Indirect. The application of the more time and resource intensive FA in these cases may suggest that there is a trend toward prioritization of intervention matched with the severity of need. The results of studies of individuals with SLD or a combination of SLD and ADHD or SLD and ED reflected single methods (FA, Direct+Indirect, and Indirect+Direct+FA) for each group of individuals, which limits the ability to compare various methods across
SLD-related disabilities. Of the eight disability categories, students without an identified disability and OHI received the widest range of assessments (3) but, due to the small numbers, the comparisons and inferences that could be made were limited. Given the aforementioned limitations that prevented the two-way ANOVA, one-way ANOVAs were conducted to compare the means of the independent groups by disability and assessment method. The independent variables were disability type and assessment where the dependent variable was the effect size. The results of the 2 one-way ANOVAs are summarized in Table 8.

Based upon the results of the one-way ANOVAs, there was no statistically significant difference for disability type or assessment method. As with the previous analysis, the assumptions for randomness and independence were met based upon review of the selected studies. Both effect sizes of assessment method and disability type reflected normal distribution although the sample size was limited. The assumptions for homogeneity of the effect sizes were met and the Levene’s tests of homogeneity were nonsignificant indicating that this assumption was met for the 2 one-way ANOVAs as well.

Table 8
Analysis of Variance for Disability Type and Method of Assessment on Effect Sizes

<table>
<thead>
<tr>
<th>Source</th>
<th>$df$</th>
<th>SS</th>
<th>MS</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disability type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between groups</td>
<td>6</td>
<td>39.45</td>
<td>6.76</td>
<td>2.21</td>
</tr>
<tr>
<td>Within groups</td>
<td>21</td>
<td>62.51</td>
<td>2.98</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>27</td>
<td>101.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assessment method</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between groups</td>
<td>3</td>
<td>5.64</td>
<td>1.88</td>
<td>0.47</td>
</tr>
<tr>
<td>Within groups</td>
<td>24</td>
<td>96.31</td>
<td>4.01</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>27</td>
<td>101.95</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
These findings suggest that there may not be a statistically significant relationship between an individual’s disability and the assessment method that effects behavioral change. Practically, the effect-size estimates do indicate that the FBA methods are effective in altering problematic behavior; however, it cannot be inferred that one method is favorable given certain student disabilities.

**Research Question 3**

The third research question addressed whether or not there was a statistical and practical difference given a particular order or sequence of conducting the FBA-based methods. Prior to this analysis, tests of independence, randomness, normality, and homogeneity were tested, and none of the assumptions were violated for the methods of assessment and the order of the methods presented. Of the 28 single cases reviewed in this analysis, there was no practical or statistically significant difference between the assessment methods overall, which is summarized in Table 9. In 21 of the cases, individuals received an FA, Indirect+Direct+ FA, or Direct+Indirect+ FA for which there is no discernible statistical or practical difference for the order of the assessment overall. There is a slight, yet nonsignificant, difference between Direct+FA and the other ordered methods, which may support the use of indirect methods as means to gain further insight into consequences maintaining the occurrence of the target behavior. The only comparison that could be calculated within a method of assessment was between Direct+ Indirect+FA ($f=2$; $ES=3.40$) and Indirect+Direct+ FA ($f=15$; $ES= 3.18$) but, because of the low numbers and the fairly comparable effective size, the contrast is not significant. As in the analysis of the first research question, the use of the method that was not FA-based, Direct+Indirect,
should be viewed with caution because the data represent a single individual and the studies were presented through nonpeer-refereed publications.

<table>
<thead>
<tr>
<th>Assessment Method</th>
<th>f</th>
<th>ES</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indirect + Direct + FA</td>
<td>15</td>
<td>3.18</td>
<td>2.02</td>
</tr>
<tr>
<td>Direct + FA</td>
<td>5</td>
<td>2.26</td>
<td>3.87</td>
</tr>
<tr>
<td>FA only</td>
<td>4</td>
<td>3.13</td>
<td>5.48</td>
</tr>
<tr>
<td>Direct + Indirect + FA</td>
<td>2</td>
<td>3.40</td>
<td>1.07</td>
</tr>
<tr>
<td>Direct + Indirect + Direct</td>
<td>1</td>
<td>5.50</td>
<td></td>
</tr>
<tr>
<td>Indirect + Direct</td>
<td>1</td>
<td>2.64</td>
<td></td>
</tr>
</tbody>
</table>

Therefore, there appears to be no statistically or practically significant difference in the effect sizes based on the order of the method used. These results suggest that there may be some flexibility or at least no apparent disadvantage in varying the order of the assessment; however, these results are inconclusive.

Research Question 4

The fourth research question involved the use of a specific framework to guide the FBA process through both the assessment and implementation phases. Furthermore, this question explored whether or not there is a difference in overall effect if a two-leveled model (Sugai, Lewis-Palmer, & Hagan-Burke, 1999) or a multiphase model (Ervin et al., 2001) was used in conducting the FBA and implementing intervention. Through this analysis of 28 individual data sets, two additional elements of the FA surfaced: (a) the length of the FA and (b) the type. Both the length and type of the FA may hold implications in bridging the research-to-practice gap.

The literature review and coding process found 24 cases utilized the multiple-level framework as compared with 4 assessments where a two-level method was
used. The results of three independent-samples $t$ tests and effect sizes for model, length, and type are presented in Table 10. Although the increased amount of time and resources required to complete an experimental manipulation such as a multielement or reversal design, 15 assessments employed experimental or quasi-experimental designs. This tendency toward more experimental methods may be linked to larger number of typical agents who are involved as the sole agent or assisting in the development and implementation of the FBAs reviewed in this meta-analysis. Given the representation of experimental analyses, it was not surprising to find that in terms of the type of FA conducted that the overwhelming majority ($f=22$) of the single-case studies utilized the more traditional method of FA that assessed for both the antecedent and consequence conditions. In contrast, 4 individuals received an FA that involved more environmentally- or antecedent-based functional assessments that were absent in the consequence condition.

The multiphase and two-level models for conducting an FA were practically significant with extremely large effect sizes, the difference between the multiphase two-level models was not statistically significant. These results should be viewed with caution as there were only 4 individuals who received a two-level FBA. In terms of practical application, the results are inconclusive; however, the average effect size for both models suggest that both yield positive behavioral change. None of the studies reviewed indicated whether there was a time or resource benefit for one model over the other.
Table 10
Results of Effect-Size Comparison by Model, Length, and Type of Functional Assessment

<table>
<thead>
<tr>
<th>Variable</th>
<th>f</th>
<th>t</th>
<th>df</th>
<th>ES</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multi- or 24</td>
<td>3.02</td>
<td>2.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-phase</td>
<td></td>
<td></td>
<td>24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two-level  4</td>
<td>3.49</td>
<td>1.68</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total      28</td>
<td>-0.45</td>
<td>26</td>
<td>22.38</td>
<td>2.81</td>
<td>1.78</td>
</tr>
<tr>
<td>Length*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td>15</td>
<td></td>
<td></td>
<td>3.41</td>
<td>2.31</td>
</tr>
<tr>
<td>FA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brief FA   11</td>
<td>2.47</td>
<td>1.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total      26</td>
<td>1.33</td>
<td></td>
<td>22.38</td>
<td>2.81</td>
<td>1.78</td>
</tr>
<tr>
<td>Type*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABC FA     22</td>
<td>2.99</td>
<td>1.94</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AB FA      4</td>
<td>3.13</td>
<td>2.34</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total      26</td>
<td>-0.13</td>
<td>24</td>
<td></td>
<td>3.06</td>
<td>2.14</td>
</tr>
</tbody>
</table>

*The variables add up to 26 as two studies did not involve an FA as part of the procedure.
Note: A Welch-Aspin test was conducted given that the assumption of homogeneity of variance was not met for Length

The effect-size difference was appreciably greater, nearly one SD, with the experimental design over the brief FA in terms of overall behavioral change; however, the results of the independent-samples t test indicated that the difference was not statistically significant. This comparison does allow for a reasonable analysis given the number of individuals in each group.

As with the model of FA, the type of FA, ABC versus AB, yielded practically significant results with extremely large, yet fairly comparable, effect sizes that should be viewed with caution as the numbers of individuals in the comparison are discrepant. Furthermore, the results of the t test found the difference between the means to be statistically nonsignificant. Following the tenets of applied behavioral analysis, the importance of the maintaining consequence and the absence of this condition in the AB type of assessment would seem to suggest that there would be a
statistical difference between the two types, which suggest that the results may be inconclusive at best.

As with the preceding effect-size assumptions for independence, normality, and randomness, these same assumptions were met in the independent-samples $t$ tests analyzed above. The assumptions for homogeneity of variance were met in all cases except for length. A significant Levene’s test, $F=9.41, p<.01$, indicated that the variances between the experimental and brief FA groups were unequal. A Welch-Aspin test was conducted given that the assumption of homogeneity of variance was not met; therefore, an independent-sample $t$ test with equal variances not assumed was used for the length comparison.

Summary

The results of literature search identified 15 studies that met the inclusion criteria for this meta-analysis that examined FBA-based assessments and interventions with students with high-incidence disabilities in public-school settings. Of those included studies, 28 individual data sets were used to calculate effect sizes to address the four research questions presented in this current research. These effect sizes ranging between small (0.14) to extremely large (7.68) were obtained using Cohen’s $d$ index. In response to the first question related to the average effect size for positive behavioral change across the various methods of assessment and disability types, the results indicated that there is a practically and significant effect for FBA-based methods and interventions. The second question involved the analysis of involved a comparison of the various methods across given disability types. Because of the limited number of cases found in the literature search, 2 one-way ANOVAs were conducted in lieu of the originally intended two-way ANOVA to compare the
effect sizes for the independent groups. The results of the one-way ANOVAs indicated that there is no statistically significant difference between the effect sizes to suggest a relationship between disability type and assessment method. The third research question was related to the order of the FBA method and whether there was a difference depending upon the order in which the assessment was conducted. The results of the analysis for assessment order revealed no practically or statistically significant difference between the effect sizes. The results of the fourth research question suggested affirmatively that the studies analyzed in this research did indeed follow a specific framework that was either multiphased or two-leveled. The findings of the three independent-samples $t$ tests indicated that there is no statistical or practical difference between the model, length, or type of FA; however, as with all of the studies analyzed within this meta-analysis, there was practical significance, positive behavioral change in the individuals, which lends credence to the mandates for FBA-based practices and positive behavioral supports.
CHAPTER V

DISCUSSION, RECOMMENDATIONS, AND CONCLUSIONS

This chapter includes a summary of the study, the significance of the study, limitations, discussion of the research questions, recommendations for practice, recommendations for future studies, and conclusions.

Summary of the Study

The purpose of the present study was to conduct a meta-analysis of single-case studies with students demonstrating pervasive and maladaptive behaviors and who were identified with high-incidence disabilities, such as emotional disturbance, learning disabilities, and other health impairment, in the public-school setting. A comprehensive search of the literature included 357 titles and abstracts related to functional behavioral assessments (FBA) and positive behavioral interventions. Only 15 of the original 357 articles and nonpublished material met the inclusion criteria for the meta-analysis. To be part of the analysis, the studies had to include a baseline and intervention phase that included at least one of three FBA methods: indirect, direct, or functional assessment (FA). Twenty-eight individual data sets reflected varying high-incidence disability types and methods of FBA. An effect size was calculated to assess behavioral change. These effect sizes were then analyzed further to address the four research questions.

The first research question concerned the average effect size for each of the various FBA-based methods and interventions for students with high-incidence disabilities, such as learning disabilities, attention deficit disorder with hyperactivity, or emotional/behavioral disorders. The second question examined which FBA methods yielded the greatest average effect sizes, and how the various methods
compared with each other across the various high-incidence disabilities. The third question addressed the difference in the effect sizes based on the order in which the methods were used, if multiple FBA methods were used. The fourth question examined whether the assessment and intervention process followed a specific framework, such as a multiple phase or two-level model, from beginning to end and, if so, was the difference in effect sizes for the interventions developed within the given framework.

The results of the meta-analysis support the use of FBA and behavior intervention plans (BIP) as sound practice in treating pervasive and maladaptive behaviors among students with high-incidence disabilities in public-school settings. The results of the one-way analysis of variances (ANOVAs) indicated that there was no statistically significant difference between the effect sizes suggesting a relationship between disability type and assessment method. Similarly, the results suggested that there was no practical or statistical effect for order of the assessment method. Although the FBA methods analyzed did follow a specific framework, there was no practically or statistically significant difference between the models and methods examined in the fourth research question; however, these findings do provide insights for practitioners.

The four research questions are discussed independently in the following sections. This chapter includes an interpretation of the results presented in the previous chapter, possible limitations with these results, and implications for practitioners as well as direction for future studies. Because FBA stems from an applied behavioral-analytic foundation that historically has been clinical in nature, the
various assessment methods are examined in an effort to bridge the research-to-practice gap by providing empirically based guidelines for practitioners.

Significance of the Study

At present, there is little guidance for practitioners as to what constitutes a proper FBA or BIP even though these practices are mandated under the Individuals with Disabilities Education Act (IDEA; 1997) and the Individuals with Disabilities Education Improvement Act (IDEIA; 2004). Educators and administrators are expected to adhere to a process of evaluation and intervention that was developed in clinical settings by researchers and their trained assistants. What little guidance has been offered by the research community relates primarily to individuals with more profound developmental delays who spend little if any part of their school-day in a general-education classroom. The current research was intended to bridge the gap between the literature and the practitioner by analyzing the accepted methods of conducting an FBA and comparing the outcomes of resulting BIPs in an effort to establish practical guidelines for school staff.

Given the legal mandates requiring the use of FBA within public schools, there is a need for school psychologists and classroom teachers to have a protocol for assessing pervasive and maladaptive behaviors in order to increase the effectiveness of behavioral intervention for students with more commonly occurring disabilities in the public-school setting. By examining the effect sizes for the difference of the baseline and intervention phases, practitioners can compare the methods and intervention outcomes when conducting an FBA, which may lead to more effective and consistently implemented BIPs.
Limitations

Although the limitations of each distinct method of FBA were detailed in the previous sections, there were a number of limitations inherent to the meta-analytic procedure. Furthermore, there were limitations unique to this study given the criteria for inclusion and the coded characteristics of the included studies.

The first limitation of a meta-analysis was retrievability bias, which invites the question of whether or not the studies included in the meta-analysis were reflective of the larger body of studies related to the subject. This limitation is akin to the analysis of the accuracy of samples reflecting the larger composition of the population. Studies included in meta-analyses generally are published works that may not be similar to nonpublished studies or theses that are more difficult to retrieve. This study utilized unpublished works; however, accessibility to these nonpublished works was limited to hand searches of local university archives and electronic resources. None of the studies referenced in the preceding review of the literature included nonpublished works.

The second issue related to the overemphasis on a single variable and the tendency to over generalize the implications of the results of this procedure. Most meta-analyses evaluate effect sizes, which critics suggest may lead to an under-representation of moderator variables. The analysis of effect sizes may yield results that are focused by identifying relevant variables that may increase the practical and statistical significance of a study such as the mean difference of a target behavior pre- and postintervention level. In other words, the purpose of a meta-analysis is to synthesize studies through the distillation of the common factors of multiple studies into a more comprehensible form, such as an effect size.
The heterogeneity of included studies in terms of variables and quality of the studies can threaten the validity and reliability of any meta-analysis (Glass, 1978). In addition to the more standard limitations such as the meta-analytic process, this study used the inclusion criteria originally employed by Wilson et al. (2003). These criteria excluded more qualitative studies and studies conducted prior to 1975. Unlike the majority of studies related to FBA and behavioral interventions that occur within a clinical setting with individuals with low-incidence disabilities, this body of research examined only individuals who were (a) school-aged, (b) attending public schools, and (c) determined to have a high-incidence disability such as a learning disability, emotional disturbance, or a speech and language impairment as their primary disabling condition. These limitations serve to narrow the focus of the study and to narrow the recommendations in terms of assessment process and intervention development that may be relevant to a larger number of students as well as practitioners.

A final threat to validity stems from Hedge’s (1981) observation that Cohen’s $d$ index may tend to overestimate slightly the effect size for an entire population when the sample size is below 20. Typically, a correction factor would be applied to address this bias; however, in this meta-analysis a correction factor was not applied as the calculations presented in the results were for individual effect sizes and not for study effect sizes. Therefore, no correction was applied to the test of homogeneity and the potential for slight bias noted. Additional limitations specific to each research question are noted further within the discussion section.
Research Question 1: Average Effect Size for Students with High-incidence Disabilities by Assessment Method

The IDEA (1997) and IDEIA (2004) mandated that school personnel conduct a FBA to address pervasive and maladaptive behaviors exhibited by students with special needs. These methods of assessment and positive behavioral interventions, although proven effective in clinical settings with severe and less frequently occurring disabilities, are required in public-school settings with students who are predominantly with high-incidence needs. The statutes do not provide specific methods, practical guidelines, or an empirical foundation to support the use of these assessments and interventions in public-school settings by practitioners. Therefore, the first research question examined the average effect size of the various FBA methods.

The effect size for the individual in the single-case studies was calculated as the difference between the intervention-phase mean and the baseline-phase mean divided by the standard deviation of the baseline phase (Busk & Serlin, 1992; Glass, 1978). Four methods, FA, Direct+Indirect, Direct+FA, and Indirect+Direct+FA, were reflected in the 15 studies included in this meta-analysis. The average effect sizes were all statistically and practically significant and ranged between 2.34 and 4.07 with a grand mean of 3.17 and an overall standard deviation of 2.14. These results suggest that FBA-based interventions can be highly effective in addressing pervasive and maladaptive behaviors demonstrated by students with high-incidence disabilities in public-school settings.

Similar results were noted by other researchers who found behavioral reductions approaching a zero occurrence rate or nearly 100% performance rate of
appropriate replacement behaviors with high-incidence disabilities in public-school settings (e.g., Broussard & Northrup, 1995; Ervin et al., 2001; Nelson & Reid, 2002). Although these studies analyzed similar FBA procedures, as those found in this study, the specific type of indirect, direct, or FA methods as illustrated in Table 3 (see Chapter 1) were rarely mentioned in detail. None of the studies provided efficacy comparison by method or effect-size calculations to investigate the magnitude of the interventions as was accomplished in this meta-analysis.

Furthermore, the synthesis studies of Ervin et al. (2002) and Nelson and Reid (2001) indicated that, although all of the studies occurred in public schools, there were very few studies that examined pervasive and maladaptive behaviors in students with high-incidence disabilities. Moreover, none of the studies reviewed by Nelson and Reid and Ervin et al. involved staff or other typical agents in the assessment and only a few typical agents participated in the implementation of the studies. Carr et al. (1999) reviewed similar factors in terms of FBA and BIP efficacy, implementation, and agents working with students with low-incidence disabilities; however, the study reflected similar findings in terms of the smaller number of cases that included the typical agent in typical settings in favor of research driven assessments and interventions in controlled, clinical settings. The Carr et al. study also found that atypical agents were most likely to utilize the more time- and resource-intensive FA methods, which led to interventions with the highest rate of success. As such, researchers involved with all three synthesis studies expressed great concern as to the practical generalizability and social validity for FBA procedures given the limited involvement of typical agents in typical settings.
In contrast, this meta-analysis found that 64% of the studies incorporated both typical and atypical agents in the FBA process and two of the unpublished works involved atypical agents for both the assessment and implementation phases. These findings suggest that there may be a shift to greater levels of practitioner involvement in conducting and implementing effective FBA procedures as evidenced by the large average effect sizes presented in this meta-analysis. The limited level of detail in most articles, theses, and papers made a critical analysis and comparison of the individual and incorporated methods difficult.

Another important limitation of this and other studies was the narrow body of literature involving students with high-incidence disabilities in public-school settings (Ervin et al., 2001; Nelson & Reid, 2002). The limited number of studies encountered in this meta-analysis is consistent with the review of Carr et al. (1999); of 266 studies, only 19 (7%) of the studies included Indirect or Direct only methods of FBA and none of the studies identified Indirect+FA specifically. As more universities, school districts, and professional organizations train psychologists, behaviorists, and teachers in FBA methods, the body of published and unpublished research may grow allowing for larger samples that may lead to further development of guidelines for practitioners in conducting the assessments and implementing the interventions.

Research Question 2: Comparison of the Various FBA Methods across Disability Types

Originally, the interaction, if any, between the independent variables (i.e., disability type and assessment method) and the dependent variable (i.e., average effect size) were to be compared through a two-way analysis of variance (ANOVA).
As stated previously, the low number of studies involving the high-incidence student population limited the breadth of this analysis to 15 single-case studies. Because of the limited number of individual cases, 2 one-way ANOVAs were conducted yielding statistically nonsignificant results. In terms of practical application, these findings may suggest that there is no statistically significant difference between the four FBA methods given different disability types. These results were consistent with Dahlstrom’s (2003) analysis of the psychometric properties of FBA procedures that found no practically or statistically significant difference between the methods of assessment.

The average effect-size statistics for the assessment methods by disability type were practically significant and provided encouraging support for the effectiveness of these various FBA methods across all disability types given the exceptionally large effect-size estimates. Overall, the most effective method of assessment, with a sufficient number of individuals, was the Indirect+Direct+FA method (ES=3.21) across the disability types. These results suggest that a comprehensive approach may be the most effective method for addressing serious behavioral events regardless of the identified disability; however, there is little research with larger samples to support or refute this assertion.

The only related study that analyzed disability type as a variable was conducted by Stage and Quiroz (1997). The dependent variable in that study was the outcome of the intervention and did not include the analysis of the assessment variable. The results of the Stage and Quiroz study indicated that the only significant, differentiated behavioral outcome for students with varying diagnoses was between individuals with emotional and conduct disorders. The individuals with emotional
disturbance demonstrated greater responsiveness to systematic intervention than their counterparts.

The absence of studies that examined the various methods of assessment and interventions across disability types on a scale larger than single-case studies suggests that future researchers may want to pursue larger synthesis studies to enhance generalizability. Matching effective methods of assessment with specific disability types may increase the efficiency of practitioners by reserving more time and resource consuming methods, such as Indirect+Direct+FA, for students with the greatest level of need.

**Research Question 3: Order of Methods**

The multiphase approach for conducting FBA involves at least one or more of the following phases: descriptive, hypothesis, experimental, and intervention procedures (Ervin et al., 2001). This synthesis study described the 4-phase model and how the phases may be linked in increasing the understanding of the function of the behavior. The recent study by Ervin et al. reported that nearly 50% of the studies incorporated interviews and direct observation methods into the FBA process; however, there was no analysis of the assessment methods relative to the behavioral outcomes that might allow for some critical evaluation of the multiphase model presented in the study. The current meta-analysis attempted to address whether there was an interaction effect for the order of method used in conducting an FBA.

The results of the analysis suggest that there was no statistically or practically significant difference for the order of the method. In terms of practical application, the practitioner may interpret these findings to mean that there is no exceptional benefit in conducting indirect assessments before direct assessment or vice-versa.
There are several studies that identify indirect methods as being effective in identifying possible antecedents, maintaining consequences, and reinforcement inventories, such as Hoff, Ervin, and Friman (ES=2.59; 2005) and Sterling-Turner, Robinson, and Wilczynski (ES=2.31; 2001). Other studies that analyzed indirect inventories such as the functional assessment inventory (FAI; Horner & Carr, 1997) and the functional assessment checklist for teachers and staff (FACTS; Borgmeier, 2003; Marche et al., 2000) reported increased effectiveness in identifying the hypothesized function of the behavior and in the resulting intervention models compared with FBAs that did not utilize the indirect method with the direct and FA methods. These findings suggest that a more flexible model that allows for the use of a variety of methods may be beneficial and that there is no appreciable difference for the order of methods used in the evaluation.

Research Question 4: FBA-based Methods Analyzed by Model, Length, and Type

Literature reviews that examined FBA variables typically identified that there were positive outcomes for the overwhelming majority of the studies reviewed and that the majority of the assessments included indirect or direct methods (e.g., Carr et al., 1999; Hanley, Iwata, & McCord, 2003; Sugai, Lewis-Palmer, et al., 1999). In the current meta-analysis, studies were disaggregated by FBA model, length, and type, and effect sizes were compared through independent-samples t tests.

The two models analyzed were the two-leveled (Sugai, Lewis-Palmer, et al., 1999) and the multiphase model (Ervin et al., 2001). There was no statistically significant difference between the model types, which may have been the result of the relatively low number of studies found with this student population that employed the two-level model. These findings suggest that practitioners may be able to conduct the
potentially less time-and-resource intensive multiphase method and if necessary may alternate to a two-level model that intensifies the assessment in response to the effectiveness of the intervention. These results, however, are inconclusive and warrant further investigation given limited sample sizes and detail of the procedural aspects of the assessment methods. Furthermore, given the disproportionate number of typical agents to atypical agents conducting and implementing the assessments, it may be beneficial to ascertain the feasibility and acceptability with practitioners utilizing either model.

To address the 10- to 23-hour time commitment in completing a traditional FA method, several studies developed abbreviated FA methods (e.g., Gresham, Watson, & Skinner, 2001; Northrup et al., 1991). For this study, traditional FA is defined loosely as \textit{functional assessment, functional analysis, functional analysis assessment, brief functional assessment, behavioral assessment, and functional communication} (Carr & Durand, 1994; Carr et al., 1999; Henley et al., 2003; Horner, 1994; Iwata et al., 1994; Northrup et al., 1991; Sasso et al., 2001). This meta-analysis examined the effect-size statistics for both the traditional and brief FA methods of assessment. Preliminary analysis of the average effect sizes found that the traditional method of FA is nearly one standard deviation higher than the brief FA. This practical difference, although apparently large in terms of effect size, is only 2 to 3 percentile points different under normal distribution. The \textit{t}-test results were viewed as reliable given the number of individuals; however, the outcomes were not statistically significant suggesting that there was no difference between the traditional and brief methods of FA. This lack of significance may be important for practitioners in that a traditional FA conducted by a trained behavioral consultant, which averages 10 to 23
hours to complete, may be shortened by conducting a brief FA of 2 to 3 hours over a period of time (Schill, Krathochwill, & Elliott, 1998). As such, the lack of a statistically significant difference between the brief FA and the experimental design is promising for practitioners given the relatively lower investment of time and resources in conducting a brief FA. The use of a two-level FBA model (Sugai, Lewis-Palmer, et al., 1999) in public-school settings might allow for more efficient and measured responses through the initial application of a brief FA that, if ineffective, could result in an experimental FA. Proponents of experimental models might argue that with truly pervasive and maladaptive behaviors the initial investment of time conducting a traditional FA would save time in the long run; however, this would presume that there are agents qualified to conduct such assessments in public schools. Therefore, if there is no difference between a brief FA that can be completed in as few as 20 two-minute trials and a traditional FA, then qualified practitioners may be confident in conducting brief FAs in addressing pervasive and maladaptive behavior. Future studies may refine the definitions of traditional and brief FA and compare the submethodologies further to evaluate them for effectiveness, duration, agent acceptability, and training methods for typical agents.

In addition to the length of the assessment and the model of the FA, this meta-analysis compared the antecedent-behavior-consequence (ABC; Hanley et al., 2003; Iwata et al., 1994) and antecedent-behavior (AB; Carr & Durand, 1985) models. Initially the AB technique, which is unusual given the lack of stimulus control in this method of assessment, was proposed as a time-saving method. The study by Carr and Durand (1985) suggested a positive reduction of the target behaviors with the AB method, which is consistent with this current analysis ($f=4$; $ES=3.13$). The research
conducted by Hanley et al. (2003) reported that 20.2% of the studies used the AB method, which is not unlike this meta-analysis where 16% of the studies used this shortened method. The results of independent-samples $t$ test were not statistically significant, and the low number of individuals reflected in the AB method makes any further inferences limited. Future studies are needed to assess the effectiveness of the AB model. Additionally, it may be useful to combine elements of the brief and AB FA to compare with traditional methods to identify efficient methods of assessment.

**Recommendations for Practice**

The results of this meta-analysis and the studies discussed in the literature review provide compelling evidence supporting the efficacy of the FBA process. These studies revealed that FBA methods of assessment are effective in assisting with the development of positive behavioral interventions that address pervasive and maladaptive behaviors. All of the FBA methods employed with students who have high-incidence disabilities analyzed in this meta-analysis resulted in very large effect-size statistics with a grand mean of 3.17. These results coupled with findings with similar student populations (e.g., Ervin et al., 2002; Nelson & Reid, 2001) suggest that practitioners can apply FBA methods in the public-school setting with a high degree of confidence as evidenced by the extremely large effect sizes.

Furthermore, the methods that consisted of FA (ES=3.13) or a combination of indirect, direct, and FA (ES=3.14) methods proved to have the highest level of practical significance. These findings suggest that practitioners may enhance the effectiveness of their interventions by using FA as a component of the assessment process; however, it may be incumbent upon these same practitioners to work with researchers in refining the definition of FA in terms that are understandable and
acceptable in the public-school system. Moreover, methods for effectively training practitioners and systems to ensure fidelity of both the assessment and implementation procedures must be developed.

The number of individuals within many of the disability categories limited the comparison of the various FBA methods across disability types. Although the results suggest that all of the FBA methods produce effective intervention plans that decreased problematic behavior, there was insufficient evidence to support the use of one method over another given a certain disability type. The more time-and-resource intensive FA methods and combinations FA and other methods were used consistently with more involved disability types, such as emotional disturbance, as evidenced by the large magnitude of the effect sizes for the various methods. The apparent matching of intensity of intervention with corresponding levels of need suggests that a two-leveled model (Ervin et al., 2001; Sugai, Lewis-Palmer, & Hagan-Burke, 1999) may be an effective solution to address the disparity between the behavioral needs within the public-school setting and the limited resources necessary to conduct these assessments. In other words, practitioners may be able to address lower intensity behaviors with less time-and-resource intensive assessment procedures, such as indirect and direct observation. If the resulting interventions are not effective, then a more comprehensive FBA can be conducted; however, further studies will be necessary to support this approach to FBA.

Furthermore, there was neither a practical nor a statistical difference between the multi-phase and two-level models of conducting an FBA as well as the assessment length between the brief and the experimental FA, which may be positive for practitioners. Because there appears to be no appreciable difference between the
two models and assessment lengths, practitioners may expend less time and personnel resources in conducting brief functional assessments within the public-school setting.

Recommendations for Future Studies

Future studies, whether, single-case or synthesis studies, may serve to bridge the gap between researchers and practitioners by increasing the level of detail specific to the assessment methods so that substantive guidelines can be developed for educators. Functional assessment should be defined procedurally so that practitioners can analyze the literature critically in an effort to replicate these procedures in the public-school setting.

Furthermore, these future studies may provide additional clarity for practical guidelines by examining the order, model, and length of assessment methods given larger sample sizes and greater levels of detail for the assessment methods. It also may be helpful for subsequent studies to address how indirect methods of assessment and direct methods can be used not only during the analytic phase but also during the evaluation of the intervention. In other words, if a hypothesis does not appear accurate based upon resulting data collected during the baseline phase, the practitioner may be wise to conduct additional indirect assessments in order to reconstruct the hypothesis. Using the FBA methods in an ad hoc manner as the intensity of the behavior necessitates and the individual responds to the intervention, may assist practitioners in establishing mastery of the various methods instead of adhering strictly to a linear model (i.e., following a specific sequence of Indirect, Direct, and then FA).

In addition, proceeding studies may include analyses of social validity and reliability particularly with the many forms of FAs conducted by practitioners and the
perceptions of the individual receiving the treatment in terms of quality of life, acceptability, practicality, and utility (cf. Kincaid, Knoster, Harrower, Shannon, & Bustamante, 2002; Reid & Nelson, 2002). The investment of time, in terms of training, evaluating, and monitoring the intervention, may make experimental and quasi-experimental procedures prohibitive in the public schools (Crone & Horner, 1999; Gresham et al., 1999, 2001; Horner & Carr, 1997).

Conclusion

This study was a preliminary analysis of the various methods of conducting FBA and implementing interventions for students with high-incidence disabilities in public-school settings who demonstrated pervasive and maladaptive behaviors. This study found that FBA-based interventions were effective in ameliorating problematic behaviors with this student population. This meta-analysis is distinct from the larger body of work related to FBA development and interventions because (a) the individuals examined in this study represent the largest group of students in the public-school system, (b) the meta-analysis used effect-size calculations for behavioral change to measure the relative power of the intervention and to compare the relevant independent variables, and (c) the present study attempted to combine variables previously not examined for interaction effect (i.e., disability type and assessment method, traditional and brief FAs, etc.). In establishing this framework for analyzing the literature, a comprehensive review of the literature was conducted. This literature review may serve as a valuable foundation for practitioners and researchers alike in summarizing not only the evolution of what has become FBA and positive behavioral interventions but also by highlighting current trends in the field. In conclusion, the results of this meta-analysis support the mandated use of FBA and
positive behavioral interventions within the public-school setting. These reviewed studies demonstrated that the assessment methods have a very high probability of effecting behavioral change positively in students with high-incidence disabilities.

As a cursory investigation, this study also identified several areas that warrant further investigation. Assuming that the current trend toward more comprehensive and detailed FBAs and interventions continues, future synthesis studies and meta-analyses may have larger samples to draw from, which in turn will undoubtedly bridge the current research-to-practice gap.
REFERENCES


* Indicates study included in the meta-analysis procedures.
APPENDIX A

ACRONYMS
## Acronyms

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB</td>
<td>Antecedent-Behavior</td>
</tr>
<tr>
<td>ABA</td>
<td>Applied Behavioral Analysis</td>
</tr>
<tr>
<td>ABC</td>
<td>Antecedent-Behavior-Consequence</td>
</tr>
<tr>
<td>BIP</td>
<td>Behavior Intervention Plan</td>
</tr>
<tr>
<td>BSP</td>
<td>Behavior Support Plan</td>
</tr>
<tr>
<td>CD</td>
<td>Conduct Disorder</td>
</tr>
<tr>
<td>EO</td>
<td>Establishing Operations</td>
</tr>
<tr>
<td>ES</td>
<td>Effect Size</td>
</tr>
<tr>
<td>ED</td>
<td>Emotional Disturbance</td>
</tr>
<tr>
<td>FA</td>
<td>Functional Assessment</td>
</tr>
<tr>
<td>FAI</td>
<td>Functional Assessment Interview</td>
</tr>
<tr>
<td>FAPE</td>
<td>Free and Appropriate Public Education</td>
</tr>
<tr>
<td>FBA</td>
<td>Functional Behavior Assessment</td>
</tr>
<tr>
<td>IDEA</td>
<td>Individuals with Disabilities Education Act (1997)</td>
</tr>
<tr>
<td>IDEIA</td>
<td>Individuals with Disabilities Education Improvement Act (2004)</td>
</tr>
<tr>
<td>IEP</td>
<td>Individualized Education Program</td>
</tr>
<tr>
<td>LRE</td>
<td>Least Restrictive Environment</td>
</tr>
<tr>
<td>MR</td>
<td>Mental Retardation</td>
</tr>
<tr>
<td>OSEP</td>
<td>Office of Special Education Programs</td>
</tr>
<tr>
<td>SIB</td>
<td>Self-injurious Behavior</td>
</tr>
<tr>
<td>SD</td>
<td>Discriminative Stimuli</td>
</tr>
<tr>
<td>NASP</td>
<td>National Association of School Psychologists</td>
</tr>
<tr>
<td>ODD</td>
<td>Oppositional Defiant Disorder</td>
</tr>
<tr>
<td>OHI</td>
<td>Other Health Impairment</td>
</tr>
<tr>
<td>PBS</td>
<td>Positive Behavioral Supports</td>
</tr>
<tr>
<td>PL</td>
<td>Public Law</td>
</tr>
<tr>
<td>SLD</td>
<td>Specific Learning Disability</td>
</tr>
<tr>
<td>SLI</td>
<td>Speech and Language Impairment</td>
</tr>
<tr>
<td>QLS</td>
<td>Quality of Life Survey</td>
</tr>
</tbody>
</table>
APPENDIX B

CODING SHEET
Coding Sheet

Rater’s name:

Title of the study:

1. Author of the study
2. Date of study
3. Single-case study design
   a. Reversal (e.g., A-B-A)
   b. Simple single-case (e.g., A-B)
   c. Alternating treatment (e.g., A-B-C-A)
   d. Multiple-baseline
4. Source type
   a. Peer-reviewed journal
   b. Dissertation
   c. Textbook
   d. Unpublished work
   e. Professional journal (not peer-reviewed)
   f. Published or unpublished case study
5. Gender of the individual receiving the treatment
   a. Male
   b. Female
   c. Not specified
6. Socioeconomic status of subject
   a. Low
   b. Middle
   c. High
   d. Not specified
7. Geographic location of study
8. Age of group, years
9. Age of group, range
10. Ethnicity of the individual
a. European American
b. Hispanic American
c. African American
d. Asian American
e. Native American
f. Other
g. Not Specified

11. Intervention type
   a. Counseling
   b. Environmental manipulation
c. Tangible reinforcement
d. Social reinforcement or socially-based
e. Punishment
f. Consequence only-no positive reinforcement
g. Removal of individual
h. Removal of undesired item or task
i. Token economy
j. Adult praise or attention
k. Other

12. Behavior
   a. Physical aggression
   b. Verbal aggression
c. Self-injurious
d. Property destruction
e. Possession of dangerous weapon
f. Possession of controlled substance
g. Inattention
h. Sensory seeking or avoiding
i. Other operationally defined behavior

13. Function of behavior
   a. Escape, protest, or avoid
b. Get or obtain

14. Personnel conducting assessment
   a. Classroom practitioner (teacher, classroom staff, etc.)
   b. School psychologist
   c. School behaviorist
   d. Other school specialist
   e. Researcher
   f. Research/ university affiliate, for example, research assistant, assisting researcher, graduate student, etc.

15. Individual supervising implementation
   a. Typical agent
      i. Classroom practitioner
      ii. School psychologist
      iii. School behaviorist
      iv. Other school specialist
      v. Parent
   b. Atypical agent
      vi. Researcher
      vii. Research/ university affiliate

16. Personnel implementing intervention
   a. Typical agent
      i. Classroom practitioner
      ii. School psychologist
      iii. School behaviorist
      iv. Other school specialist
      v. Parent
   b. Atypical agent
      i. Researcher
      ii. Research/ university affiliate

17. Assessment model
   a. Multiphase model
   b. Two-level
c. Other model
d. None evidenced

18. Rated intensity of service
   a. Very poor
   b. Poor
   c. Average
   d. Good
   e. Very good
   f. Not reported by researcher

19. Rated implementation quality
   a. Very poor
   b. Poor
   c. Average
   d. Good
   e. Very good
   f. Not reported by researcher

20. Primary disability
   a. Specific learning disability
   b. Speech and language impairment
   c. Orthopedically impaired
   d. Other health impaired
   e. Autism
   f. Mental retardation
   g. Emotional disturbance
   h. Visually impaired
   i. Blind
   j. Deaf/Hearing impairment
   k. Visually and hearing impaired
   l. Multiple-disabilities
   m. Section 504 eligibility only

21. Method of FBA
   a. Direct (D) only
b. Indirect (I) only
c. Functional assessment (FA) only
d. D+I
e. I+D
f. D+FA
g. FA+D
h. I+FA
i. FA+I
j. I+D+FA
k. D+I+FA
l. FA+D+I
m. FA+I+D
n. I+FA+D
o. D+FA+I

22. Rated fidelity in implementation
   a. Very poor
   b. Poor
   c. Average
   d. Good
   e. Very good
   f. Not reported by researcher

23. Level of post intervention follow-up
   a. No follow-up reported
   b. 0-3 months
   c. 3-6 months
   d. More than 6 months

24. Agent conducting follow-up
   a. Typical agent
      i. Classroom practitioner (teacher, classroom staff, etc.)
      ii. School psychologist
      iii. School behaviorist
      iv. Other school specialist
v. Parent

b. Atypical agent
   i. Researcher
   ii. Research/university affiliate

25. Time between measures
26. Baseline phase mean
27. Intervention phase mean
28. Standard deviation
APPENDIX C

STUDIES INCLUDED IN THE META-ANALYSIS
# Studies Included in the Meta-analysis

<table>
<thead>
<tr>
<th>Study coded</th>
<th>Author (Year)</th>
<th>Source</th>
<th>( f )</th>
<th>ES</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Dunlap, DePreczei, Clarke, Wilson, Wright, White, &amp; Gomez (1994).</td>
<td>Journal of Applied Behavior Analysis (JABA)</td>
<td>3</td>
<td>2.65</td>
<td>23.21</td>
</tr>
<tr>
<td>3</td>
<td>Cooper, Wacker, Thursby, Plagmann, Harding, Millard, &amp; Derby (1992).</td>
<td>JABA</td>
<td>2</td>
<td>3.40</td>
<td>1.00</td>
</tr>
<tr>
<td>4</td>
<td>Kern, Childs, Dunlap, Clarke, &amp; Falk (1994)</td>
<td>JABA</td>
<td>1</td>
<td>2.81</td>
<td>12.79</td>
</tr>
<tr>
<td>5</td>
<td>Ellis, &amp; Magee, (1999)</td>
<td>Education and Treatment of Children (ETC)</td>
<td>1</td>
<td>1.32</td>
<td>20.69</td>
</tr>
<tr>
<td>6</td>
<td>Musser, Bray, Kehle, &amp; Jenson (2001)</td>
<td>School Psychology Review (SPR)</td>
<td>3</td>
<td>6.60</td>
<td>3.96</td>
</tr>
<tr>
<td>7</td>
<td>Strichtler, Sasso, &amp; Jolivette (2004)</td>
<td>Journal of Positive Behavior Intervention (JPBI)</td>
<td>1</td>
<td>2.01</td>
<td>23.59</td>
</tr>
<tr>
<td>8</td>
<td>Meyer (1999)</td>
<td>JABA</td>
<td>4</td>
<td>3.13</td>
<td>10.08</td>
</tr>
<tr>
<td>9</td>
<td>Radford, &amp; Ervin (2002)</td>
<td>JPBI</td>
<td>1</td>
<td>3.09</td>
<td>14.32</td>
</tr>
<tr>
<td>10</td>
<td>Packenham, Shute, &amp; Reid (2004)</td>
<td>ETC</td>
<td>2</td>
<td>1.73</td>
<td>12.41</td>
</tr>
<tr>
<td>11</td>
<td>Sterling-Turner, Robinson, &amp; Wilczynski (2001)</td>
<td>SPR</td>
<td>1</td>
<td>2.31</td>
<td>14.43</td>
</tr>
<tr>
<td>13</td>
<td>Wright-Gallo (2005)</td>
<td>Master’s thesis</td>
<td>1</td>
<td>2.64</td>
<td>22.34</td>
</tr>
<tr>
<td>14</td>
<td>Daniels (2002)</td>
<td>Doctoral dissertation</td>
<td>3</td>
<td>1.64</td>
<td>11.65</td>
</tr>
<tr>
<td>15</td>
<td>Positive Environment Network of Trainers (2003, February)</td>
<td>Technical publication</td>
<td>1</td>
<td>5.50</td>
<td>10.35</td>
</tr>
</tbody>
</table>

*Note:* Frequency reflects the number of individuals who met the inclusion criteria for this meta-analysis (see Chapter 3). Therefore, not all individual cases within a given study were included in this analysis.