Functional Behavioral Assessment-Based Interventions for Students with or At Risk for Emotional and/or Behavioral Disorders in School: A Hierarchical Linear Modeling Meta-Analysis

Nicholas A. Gage
University of Connecticut

Timothy J. Lewis
University of Missouri

Janine P. Stichter
University of Missouri

ABSTRACT: Of the myriad practices currently utilized for students with disabilities, particularly students with or at risk for emotional and/or behavioral disorder (EBD), functional behavior assessment (FBA) is a practice with an emerging solid research base. However, the FBA research base relies on single-subject design (SSD) and synthesis has relied on literature review or analyses using nonparametric effect size calculations. This study was designed to examine the omnibus effect that FBA-based interventions have on problem behaviors for students with or at risk for EBD in schools using a hierarchical linear modeling meta-analytic approach to SSD synthesis. Based on a sample of 69 FBA studies, 146 subjects, and 206 outcome graphs, results indicated that, overall, FBA-based interventions reduced problem behavior by an average of 70.5% and that the procedure was effective across all student characteristics. Differences of effectiveness were evident between functional analysis and descriptive assessment procedures. Findings of this study suggest that FBA-based interventions for students with or at risk for EBD are an effective approach for the reduction of problem behaviors.

■ Since the 1997 reauthorization of the Individuals with Disabilities Education Act, schools have been required to conduct functional behavioral assessments (FBA) to develop an intervention when a student with an individual education plan is suspended beyond 10 days or placed in an interim alternative educational setting (Yell, 2006). An FBA is broadly defined as a set of methods that collect information to identify environmental events, antecedents and consequences, that predict or maintain problem behaviors. An FBA attempts to identify the "function" of a challenging behavior to develop an intervention to reduce the occurrence of that behavior and increase the occurrence of an alternative replacement behavior that results in the same or similar outcomes (Conroy & Stichter, 2003; Gresham, Watson, & Skinner, 2001; Scott & Kamps, 2007).

The goal of the FBA process is the development and implementation of an intervention (Neilsen & McEvoy, 2004). Dunlap, Kern-Dunlap, Clarke, and Robbins (1991) suggest that interventions matched to behavioral function involve three general strategies: (1) altering the antecedents; (2) establishing or strengthening the response–reinforcer relationship for alternative behaviors; and (3) weakening the maintaining response–reinforcer relationship. Evidence of effectiveness exists for function-based interventions for each of the three strategies (Christensen, Young, & Marchant, 2007; Edwards, Magee, & Ellis, 2002; Ellis & Magee, 1999; Grandy & Peck, 1997; Hoff, Ervin, & Friman, 2005; Lane et al., 2007; Liaupsin, Umbreit, Ferro, Urso, & Upreti, 2006; Moore et al., 2002; Park & Scott, 2009; Stahr, Cushing, Lane, & Fox, 2006; Trussell, Lewis, & Stichter, 2008; Umbreit, 1995; Umbreit & Blair, 1997). What has not been fully delineated is the effectiveness of the process within school environments for students with or at risk for emotional and/or behavioral disorders (EBD; Sasso, Conroy, Stichter, & Fox, 2001; Scott & Kamps, 2007). Further, although there is

a growing database of FBA research in schools led by school staff (Bessette & Wills, 2007; Kamps, Wendland, & Culpepper, 2006; Maag & Larson, 2004; Skinner, Veerkamp, Kamps, & Andra, 2009; Wallace, Doney, Mintz-Resudek, & Tarbox, 2004; Watson, Ray, Turner, & Logan, 1999), the field has yet to develop either specific guidelines or agreement on exactly what an FBA must include (e.g., qualification of assessment agent, mandatory standardized forms, age of student, behavior topography, etc.), what essential features are necessary for effective assessment-to-intervention development (e.g., functional analysis, intervention agent, settings), and whether different types of interventions, such as stimulus and reinforcement-based interventions or punishment and extinction procedures (Carr et al., 1999), are more effective in schools for students with or at risk for EBD.

FBA, Literature Reviews, and Meta-Analysis

To examine the lack of agreement of FBA procedures for students with or at risk for EBD, authors have synthesized existing research, providing recommendations for research and practice (Fox & Davis, 2005; Gresham et al., 2001; Lane, Kalberg, & Shepcaro, 2009; Lane et al., 1999; Nelson, Roberts, Mathur, & Rutherford, 1999; Weber, Killu, Derby, & Barretto, 2005). It is evident from reviews to date (Carr et al., 1999; Solnick & Ardoin, 2010; Wood, Blair, & Ferro, 2009) that there is a discrepancy between the amount of research on functional analysis versus descriptive assessment, who is conducting the assessments, where both functional analysis and interventions are being conducted, and under what circumstances the intervention results are most effective. These reviews highlight the large and growing database of studies, but none have identified whether FBA-based effectiveness is related to the key features of the FBA process, including assessment and intervention features, such as who conducts the assessment and intervention, how the assessment is conducted (descriptive or experimental), where the assessment and intervention occur, and student features, such as age and disability.

In order to provide further empirical evidence of FBA-based interventions, meta-analyses of the primarily single-subject design (SSD) literature base are warranted. In addition to the synthesis of overall findings and differential effects based on student and assessment characteristics, research into differential effects of functional analysis, the experimental manipulation of antecedents or consequences, and descriptive assessment—FBA process without experimental manipulation—would provide the field with evidence of the need, or lack thereof, for experimental analysis in FBA in schools. Further, FBA-based research has been conducted across multiple environments (classrooms, playgrounds, cafeterias) and contexts (schools and clinics); evidence further validating the procedure within each context and environment is needed. Overall, a meta-analysis of the FBA literature base could (1) contribute to the debate of its efficacy and effectiveness in schools; (2) identify differences in intervention effectiveness based on assessment and intervention features and student features; and (3) potentially provide further empirical evidence that the procedure is an evidence-based practice.

To date, three meta-analyses have been conducted on FBA-based intervention research. Marquis et al. (2000) conducted a meta-analysis of the research synthesis by Carr et al. (1999) on positive behavior supports for individuals with developmental disabilities. Using the standard mean difference of all data (SMD_{All}; the mean of the baseline data points minus the mean of the intervention data points and divided by the standard deviation of the baseline data points) as the effect size, the study found that degree of cognitive impairment was a significant predictor of effect size, with greater effects for individuals with less cognitive impairment. Additionally, interventions that utilized typical agents in typical settings had significantly greater effect sizes than those with atypical agents in atypical settings. These findings indicate that for individuals with cognitive impairment or autism or both, FBA-based interventions are more effective when the individual has less cognitive impairment and the intervention is conducted by intervention agents who are most familiar with the subject in their natural environments. Assessment procedures (descriptive and experimental) were aggregated together, so nuanced differences between them were not analyzed.

Gresham, McIntyre, Olson-Tinker, Dolstra, McLaighlin, and Van (2004) examined school-based behavior analytic FBA and PBS (Positive Behavioral Supports) research published in *Journal of Applied Behavior Analysis*. They identified 150 studies and, using the

Faith, Allison, and Gorma (1997) *d* statistic, which is calculated the same as SMD_{All}, and percentage of nonoverlapping data (PND; Scruggs, Mastropieri, & Casto, 1987), they found significantly larger effect sizes when the study employed experimental analyses, such as functional analysis, compared with descriptive assessments alone. However, they found that non-FBA-based interventions had the largest effect size across both effect size calculations, yet cautioned that there was significant variability, particularly in the non-FBA effect sizes.

Recently, Goh and Bambara (in press) conducted a meta-analysis of individualized positive behavior supports (IPBS) in school settings. They identified 83 studies with 145 participants and, using PND, found IPBS to be effective in schools (PND = 88%). In addition to the identification of overall effectiveness, Goh and Bambara examined differences resulting from student characteristics and assessment and intervention characteristics and found moderate to large effects across all variables. Although they included studies with students with developmental disabilities (cognitive impairment and autism), they found IPBS interventions to be effective for students with EBD and students at risk. Their findings extend those of other reviews by examining FBA-based interventions in schools. However, the synthesis procedure relied on PND for identification of effectiveness, which (1) is nonparametric and (2) does not meet SSD meta-analysis characteristics recommended by Wolery, Busick, Reichow, and Barton (2010).

Hierarchical Linear Modeling Meta-Analysis

One promising approach in SSD meta-analysis that has the potential to meet criteria developed by Wolery et al. (2010) is the application of hierarchical linear modeling (HLM) procedures (Jenson, Clark, Kircher, & Kristjansson, 2007; Nagler, Rindskopf, & Shadish, 2008; Van den Noorgate & Onghena, 2007, 2008). This approach aggregates multiple studies together at the subject level (thus increasing sample size exponentially), calculates robust *t*-ratios using maximum likelihood estimation for significance testing, and calculates an effect size that addresses serial dependence.

Potential advantages of HLM for meta-analysis of SSD using each subject's data as displayed in SSD graphs include the automatic calculation of random effect models to account for autocorrelation and the ability to simultaneously analyze subject data even when the subjects have different numbers of observations over time (Raudenbush & Bryk, 2002). The presence of autocorrelation—the correlation of the errors for observations—underestimates the standard error in conventional linear models and makes the significance test misleading. However, random effects, which are extra values associated with each subject and do not have substantive meaning, can account for and correct autocorrelation in HLM models, assuming a large sample size (Guo & Hussey, 1999). Unlike other approaches, including PND, SMD_{All}, or other effect sizes for SSD (for a comprehensive review, see Maggin, Swaminathan, Rogers, O'Keefe, Sugai, & Horner, 2011), an HLM meta-analysis does not calculate an individual effect size but rather utilizes all of the data available for the analysis.

Purpose

This study extends the work of previous reviews of FBA-based intervention research by examining only studies conducted in schools with students with or at risk for EBD. In addition, the synthesis procedure utilized HLM meta-analysis, a parametric approach that models intervention effects along with moderating variables and results in both effect sizes and statistical significance values for mean shift, slope, and variability in SSD research (Nagler et al., 2008; Van den Noorgate & Onghena, 2003). Specific research questions were (1) how effective FBA-based interventions are for students with or at risk for EBD in schools, (2) whether student, assessment and intervention, and study (e.g., study quality) characteristics impact the results, and (3) how the results of this analysis compare with the results of earlier meta-analyses.

Method

Study Identification and Selection

Beginning with Iwata, Dorsey, Slifer, Bauman, & Richman's (1982/1994) seminal study, published, peer-reviewed treatment studies utilizing FBA procedures to develop interventions for students with or at risk for EBD were identified through a two-stage review method.

Stage 1

Using the key terms "functional behavior assessment," "function based intervention," "structural analysis," "behavior intervention," "functional analysis," or a combination of terms, electronic searches of PsycLit, ERIC, and Google Scholar databases were conducted. In addition, hand searches of primary journals in the field (Campbell, 2003) were conducted to identify FBA-based intervention studies. Journals searched included *Behavioral Disorders, Journal of Emotional and Behavioral Disorders, Journal of Positive Behavior Interventions, Education and Treatment of Children,* and *Journal of Applied Behavior Analysis.* Finally, ancestral reviews of identified articles were conducted to ensure that all FBA articles within the literature were identified.

Overall, 434 articles were identified during the stage 1 literature search.

Stage 2

Articles were then selected for inclusion based on the following criteria.

1. Articles were excluded if the subjects were cognitively impaired or had an autism spectrum disorder (ASD), including autism and pervasive developmental disorder. Students with EBD, attention deficit–hyperactivity disorder (ADHD), or learning disabilities or not identified with any disability were included in the study as long as the target behavior was not academic performance. The inclusion criterion implies that a student without cognitive impairment or ASD qualifies as at risk for EBD if an FBA is conducted in an attempt to change the student's behavior. If a study had at least one student meeting the criteria, the study was included, but only the data from the student meeting the criteria were included in the results.
2. Only articles utilizing SSD were included, and each study must have included the following: (1) a graphic display for each student; (2) a replicable baseline, defined as sufficiently precise information for other researchers to replicate (Horner et al., 2005), and at least one treatment phase present in each graphic display; and (3) explicitly described independent and dependent variables and their measurement dimensions. These three features were necessary for the analytic procedures because the designs needed to be similar to aggregate the data (i.e., a study without a baseline is not comparable with a study that uses a baseline because there is no control with which to compare intervention effects).
3. Studies must have explicitly stated the procedures used to identify and confirm the hypotheses developed in the FBA assessment (e.g., descriptive assessment or functional analysis). Descriptive assessment was defined as the development of an intervention based on hypotheses of function from indirect (rating scales and interviews) and direct (behavior observations) assessment methods. Functional analysis was defined as the experimental manipulation of antecedents and consequences using analog conditions to systematically control and assess the function of a behavior to develop an intervention (Iwata et al., 1982/1994).

Of the 434 identified articles, 69 met the inclusion criteria. Of the articles, 28% (n = 19) were from *Behavioral Disorders,* 26% (n = 18) were from *Education and Treatment of Children,* 13% (n = 9) were from *Journal of Applied Behavior Analysis,* and 9% (n = 6) were from *Journal of Positive Behavior Support.* The percentage of articles that utilized functional analysis procedures (59%) was slightly greater than those that utilized descriptive assessment procedures (41%).

Coding Frame

Each individual study was coded for analysis. Student-level, assessment and intervention-level, and study-level characteristics in *Table 1* were individually coded for each student and each of their outcomes (i.e., engagement and physical aggression data) for analysis.

Graphic Data Extraction

Outcome data were collected via the graphic display of data in each study using Biosoft's Ungraph 5.0. Recent research has validated Ungraph as a reliable and valid method of data extraction (Shadish et al., 2009), and the program is recommended in a procedural manual for conducting single-subject meta-analysis (Nagler et al., 2008).

A total of 5,189 data points were individually coded, with 2,009 baseline data points and 2,977 intervention data points. Eight studies compared the effectiveness of FBA-based

TABLE 1
Coding Frame

Student-Level Characteristics	Assessment and Intervention Characteristics	Study-Level Characteristics
Student age: Age was recorded in years. Studies that provide age in year and month were rounded to the nearest year and studies that provide only grade were recorded as the typical age of students at the beginning of that grade (e.g., first grade would be recorded as 6 years). *Student gender*: Gender was recorded as a dichotomous male–female variable. *Disability*: The primary disability of each student was recorded. If a student had a co-morbid disability, it was also recorded.	*Functional analysis*: Recorded whether the study authors conducted a functional analysis, including structural analysis. If the study did not conduct an experimental manipulation of either the antecedent or the consequence within the assessment phase, it was coded as descriptive assessment. *Experimental manipulation setting*: The environment in which the experimental manipulation was implemented (e.g., clinic, special education classroom, general education classroom), if applicable. *Intervention setting*: The environment in which the intervention was implemented (e.g., clinic, special education classroom, general education classroom). *Assessment agent*: Who conducted the assessment phases (e.g., teacher or researcher). *Intervention agent*: Who conducted the intervention phase of the study (e.g., teacher, researcher, graduate student, or paraprofessional). *Behavior topography*: Participant's behavior topography was coded into one of the following categorical variables: physical aggression, verbal aggression, noncompliance, social withdrawal, initiations, engagement, or multiple behaviors. Although many studies report multiple behavior topographies for individual students, the primary targeted behavior was coded. Operational definitions were developed and used in IOA calculations. *Behavioral intervention technique*: The type of intervention technique utilized. Four distinct categorical codes, based on Campbell (2003), were used: (1) punishment, including any intervention designed to reduce the likelihood of a behavior, such as overcorrection or timeout; (2) positive procedures, including any intervention that presented a stimulus to increase the likelihood of a behavior, such as positive reinforcement; (3) a combination of punishment and positive procedures; and (4) extinction or withholding reinforcement.	*Quality indicator score*: Score based on Horner et al. (2005) quality indicators. Horner et al. outlined 21 individual indicators within 7 categories and each study was assessed for how well it met the quality indicators by identifying the number of indicators the study met. The quality indicator scoring procedure for each article was conducted using a standardized form developed by Hudson, Lewis, Stichter, and Johnson (2011) that defined each indicator in two sections: essential and desired features of SSD. The form was used for both data collection and inter-rater agreement. Because essential indicators are considered mandatory components of SSD (Horner et al., 2005), they were weighted in the final quality indicator score following the recommendations from Lane et al. (2009). Each essential quality indicator was given a numeric value of 2, while desirable quality indicators were given a numeric value of 1. The range of possible scores for the quality indicator variable was 0 to 23. *Autocorrelation*: Lag-1 autocorrelation was calculated for each graphic display.

Note: Operational definitions used for inter-rater reliability can be obtained directly from the first author.
IOA: Inter-observer Agreement.

interventions with that of other interventions, including Check-In/Check-Out (n = 2), First Step to Success (n = 1), and non-function-based interventions (n = 5), which accounted for only 203 data points. Due to the limited amount of data for these non-FBA-based interventions, these data points were removed from the study because they were not directly comparable and could not be modeled as a distinct comparison phase in the analysis.

A concern in the synthesis of single-subject research is the variability of measurement dimensions used for the dependent variable (DV) and the need for a single dimension across studies. Additionally, the direction of the intervention effect is contingent on the outcome variable (e.g., increase prosocial behavior or decrease problem behavior). The outcome variable reported here was scaled as percentage and the data were recoded so intervention effect should have a negative slope. Of all the studies, 88% used percentage as the DV, 7% used frequency, and 5% used rate. All frequency and rates were recalculated into percentage by dividing the session score for each data point by the total

number possible (largest number printed) on the ordinate axis. After transformation, graphic displays were examined to confirm equality of magnitude change (i.e., effect of intervention equal across scaling).The slope transformation relied on a reverse coding procedure using the following formula on all data points in graphs with positive slopes (e.g., increasing use of a positive behavior):

$$\text{new value} = (\text{scale minimum} + \text{scale maximum}) - \text{old value}$$

Recoding of both the scale and slope direction were conducted using Microsoft Excel.

Additionally, all data and results are reported at the outcome level, thereby using all the data available for the analysis. The outcome analysis is built on the logic of SSD that the focus is a behavior, not a subject; therefore, the analysis of an aggregate behavior change is in line with the assumptions of SSD because the key is the behavior change as a result of the independent variable (Kazdin, 1982). The aggregation of multiple outcome effect sizes for a single subject for group designs in traditional meta-analysis has statistical validity and is used in practice (Borenstein, Hedges, Higgins, & Rothstein, 2009).

Statistical Procedures

The first step in the analytic process was a descriptive analysis of student and study characteristics. All descriptive characteristics were analyzed using SPSS 16.

Single-Subject Effect Sizes

PND was calculated for each individual student's graph by computing the percentage of data points in the treatment phase that exceed the most extreme baseline data point that is in the intended treatment direction as outlined by Scruggs et al. (1987). PND was chosen because it is among the oldest effect size calculations in SSD and is the most frequently used in special education (Wolery et al., 2010) and for result comparison with Goh and Bambara (in press). Additionally, the SMD_{All} was calculated for comparison with Marquis et al. (2000) and Gresham et al. (2004). The SMD_{All} (Busk & Serlin, 1992) is similar to Cohen's *d*, where the mean of the baseline phase is subtracted from the mean of the treatment phase and divided by the standard deviation of the baseline phase. The scores were averaged across all studies for a single effect size and compared with the results of the HLM meta-analysis.

HLM Meta-Analysis

The HLM meta-analysis was conducted in multiple steps using the HLM 7.0 software program (Raudenbush & Bryk, 2010). A series of models was developed examining the tenability of the data using the HLM meta-analytic approach and the influence of phase ordering and phase by order by session interactions. In the model used, the dependent variable score was the sum of four parts: (1) the intercept, or the mean score at baseline; (2) a term accounting for the rate of change in count with sessions; (3) a term accounting for the rate of change in count with phase changes; and (4) an interaction term allowing the rate of change in count to differ across phases. The key component values in this analysis were the β coefficients and *p*-values for the average rate of change as students switched from baseline to treatment phases for all students, an error term that allows each student's DV to vary around the grand mean, and the β coefficient and *p*-values for the average change in session effect (time slope) as a student switched from baseline to treatment phase for all students (Nagler et al., 2008). In this analysis, the initial status is the mean at baseline, the baseline slope is the rate of increase or decrease for each data point in baseline, the intervention effect is the mean level change between baseline and intervention (mean shift), and the intervention slope is the rate of increase or decrease for each data point in the intervention.

To examine the impact of moderating variables, a series of models was developed to identify the effect of student-level, assessment and intervention-level, and study-level characteristics from the coding frame in *Table 1* on the fixed effects. These variables were included because they are substantively important from the studies (Kavale, 1984) and because the FBA literature base has not identified whether student characteristics such as age, disability, or behavior topography impact intervention results (Carr et al., 1999). Each of the variables described in the coding frame extracted for each of the students were explored. Taken together, the analytic procedure provides both inferential interpretation of FBA-based interventions and an effect size for the procedure (see Appendix A).

Inter-Rater Reliability

Inter-rater reliability was assessed using a random sample of 57% of the included articles (n = 39) and was calculated by dividing the total number of agreements by the total number of agreements plus disagreements for each item and averaged across all items. The first author and two graduate students coded the articles for all moderating variables (e.g., age, disability, behavior topography, etc.), and inter-rater reliability was 92% (range: 82–100%) across all variables. The quality indicator instrument inter-rater reliability was assessed separately by the first author and two graduate students to ensure reliability of the overall score. All 69 quality indicator instruments were dual coded, and a final agreement across all of the indicators was 87.3%.

Results

Student-Level Characteristics

A total of 146 students were identified within the 69 included studies. *Table 2* provides a comprehensive breakdown of student and study characteristics by article. Behavioral topography at the student level was also coded and served as the DV. Behavior topography was variable, with most students exhibiting multiple behaviors of concern (62.8%). Overall, off-task, out-of-seat behavior, inappropriate verbalizations (including swearing, arguing, and yelling), and physical aggression were prevalent. Three general categories broadly describe the DVs measured within the studies. The most prevalent DV was an aggregate of problem behaviors (38%), typically called "disruptive behavior" or "challenging behavior." "Engagement" (34%) was the second most prevalent, and an "other" category that included all single behaviors (28%), such as yelling, aggression, and talk-outs, was the third most prevalent. Additionally, intervention type was coded. The majority of studies (90.5%) utilized positive procedures in the intervention, while 5.4% used a combination of punishment, extinction, or positive procedures.

Quality Indicator Scores

The mean quality indicator score across all 69 articles was 16.1, with a standard deviation of 3.4, a range of 7 to 21, and a median of 17. Although there was some variability between the articles, overall, the quality of the FBA-based intervention research was moderately high. The median possible value from the calculation was 11.5, yet approximately 90% (n = 61) of the articles scored above that. None of the articles scored the maximum possible value (23). Examination of just the essential quality indicators revealed a mean score of 13.2, a standard deviation of 2.5, a range of 6 to 16, and a median of 14. These results suggest that the studies are meeting most of the essential quality indicators.

Nonparametric Single-Subject Effect Sizes

Effect sizes were calculated for all data at both the student level and the outcome level. A total of 412 individual effect sizes were calculated. However, two limitations to the effect sizes were identified during their calculations: (1) PND scores biased results if one baseline point was at the minimum allowable value for the outcome (specifically, if the target behavior was "problem behavior," a baseline value of 0 resulted in an overall PND score of 0), and (2) SMD_{All} could not be calculated for graphs with fewer than three data points during baseline or if there was no deviation in the baseline values. Overall, 25 outcomes had a PND value of 0 and 12 outcomes had either fewer than three baseline data points or a standard deviation value of 0.

Table 3 provides descriptive statistics of the effect sizes. Pearson's product moment correlations (r) were used to examine the association between the PND and SMD_{All} results. The correlation value was $r = .432$ and $r = .477$ at the outcome level. Both correlation coefficients were significant at $p < .01$.

HLM Meta-Analysis

First, an empty model was examined to calculate an intraclass correlation coefficient (ICC). The ICC value was .217, meaning that 21.7% of the variance in the dependent variable was between students and 78.3% of variance was within each student's DV score. This indicates that since the majority of variability within the data was within each student's DV, the pattern of dependent variable change between students was similar.

TABLE 2
Included Articles and Coding Frame Characteristics

Articles	Student	Age	Gender	Disability	FA	Assessment Setting	Intervention Setting	Experimental Design	Assessment Agent	Intervention Agent
Bessette & Wills (2007)	1	8	Male	EBD	FA	SPED	SPED	AB	Para	Para
Blair et al. (2010)	1	3	Male	Language disorder	Descript	N/A	Preschool	MB	Researcher	Teacher
	2	4	Male	ADHD	Descript	N/A	Preschool	MB	Researcher	Teacher
Blair et al. (1999)	1	5	Male	EBD	FA	Preschool	Preschool	MB	Researcher	Teacher
	2	5	Female	EBD	FA	Preschool	Preschool	MB	Researcher	Teacher
	3	5	Male	EBD	FA	Preschool	Preschool	MB	Researcher	Teacher
	4	5	Male	EBD	FA	Preschool	Preschool	MB	Researcher	Teacher
Boyjian et al. (2001)	1	4	Male	ADHD	FA	Preschool	Preschool	CR	Researcher	Combination
	2	4	Male	ADHD	FA	Preschool	Preschool	CR	Researcher	Combination
	3	5	Male	ADHD	FA	Preschool	Preschool	CR	Researcher	Combination
Broussard & Northrup (1995)	1	8	Male	At risk	FA	Gen ed	Gen ed	CR	Researcher	Researcher
	2	6	Male	At risk	FA	Gen ed	Gen ed	CR	Researcher	Researcher
	3	6	Male	ADHD	FA	Gen ed	Gen ed	CR	Researcher	Researcher
Broussard & Northrup (1997)	1	9	Male	ADHD	FA	Lab school	Gen ed	AB	Researcher	Researcher
	2	7	Male	ADHD	FA	Gen ed	Gen ed	AB	Researcher	Researcher
	3	7	Male	At risk	FA	Gen ed	Gen ed	AB	Researcher	Researcher
	4	6	Male	At risk	FA	Gen ed	Gen ed	AB	Researcher	Researcher
Campbell & Anderson (2008)	1	10	Male	At risk	Descript	N/A	Gen ed	CR	Researcher	Team
	2	10	Male	At risk	Descript	N/A	Gen ed	CR	Researcher	Team
Carter & Horner (2009)	1	6	Male	At risk	Descript	Gen ed	Gen ed	MB	Researcher	Team
	2	7	Male	At risk	Descript	N/A	Gen ed	MB	Researcher	Team
	3	5	Male	At risk	Descript	N/A	Gen ed	MB	Researcher	Team
Christensen et al. (2004)	1	8	Male	At risk	Descript	N/A	Gen ed	ABAB	Teacher	Teacher
	2	8	Male	At risk	Descript	N/A	Gen ed	ABAB	Teacher	Teacher
Christensen et al. (2007)	1	8	Male	LD	Descript	N/A	Gen ed	ABAB	Teacher	Teacher
Clarke et al. (1995)	1	5	Male	EBD	Descript	N/A	SPED	ABAB	Researcher	Teacher
	2	11	Male	EBD	Descript	N/A	SPED	ABAB	Researcher	Teacher
	3	11	Male	EBD	Descript	N/A	SPED	ABAB	Researcher	Teacher
Cooper et al. (1992)	1	9	Male	EBD	FA	SPED	SPED	MB	Combination	Teacher
	2	8	Male	EBD	FA	SPED	SPED	MB	Combination	Teacher

TABLE 2
Continued

Articles	Student	Age	Gender	Disability	FA	Assessment Setting	Intervention Setting	Experimental Design	Assessment Agent	Intervention Agent
DePaepe et al. (1996)	1	12	Male	EBD	Descript	N/A	SPED	ABAB	Researcher	Teacher
	2	9	Male	EBD	Descript	N/A	Isolated	ABAB	Researcher	Teacher
Dufrene et al. (2007)	1	5	Male	At risk	FA	Preschool	Preschool	ABAB	Researcher	Combination
	2	5	Male	At risk	FA	Preschool	Preschool	ABAB	Researcher	Combination
	3	5	Female	At risk	FA	Preschool	Preschool	ABAB	Researcher	Combination
Dunlap et al. (1994)	1	11	Male	EBD	Descript	N/A	SPED	ABAB	Researcher	Teacher
	2	11	Male	EBD	Descript	SPED	SPED	ABAB	Researcher	Teacher
	3	5	Male	EBD	Descript	SPED	SPED	ABAB	Researcher	Researcher
Dunlap et al. (1996)	1	7	Male	EBD	Descript	N/A	SPED	ABAB	Researcher	Teacher
	2	9	Female	EBD	Descript	N/A	SPED	ABAB	Researcher	Teacher
	3	7	Female	EBD	Descript	N/A	SPED	ABAB	Researcher	Teacher
Durand & Carr (1992)	1	5	Male	ADHD	FA	Isolated	Isolated	MB	Researcher	Researcher
	2	4	Male	ADHD	FA	Isolated	Isolated	MB	Researcher	Researcher
	3	3	Male	Language disorder	FA	Isolated	Isolated	MB	Researcher	Researcher
	4	5	Male	Language disorder	FA	Isolated	Isolated	MB	Researcher	Researcher
	5	3	Male	Language disorder	FA	Isolated	Isolated	MB	Researcher	Researcher
	6	5	Male	Language disorder	FA	Isolated	Isolated	MB	Researcher	Researcher
Edwards et al. (2002)	1	10	Male	EBD	FA	Isolated	Multiple settings	AB	Researcher	Team
Ellis & Magee (1999)	1	10	Male	ADHD	FA	Isolated	Multiple settings	AB	Researcher	Researcher
Ervin et al. (1998)	1	13	Male	ADHD	FA	SPED	SPED	ABAB	Combination	Teacher
	2	14	Male	ADHD	FA	SPED	SPED	MB	Combination	Teacher
Fairbanks et al. (2007)	1	7	Male	At risk	Descript	Gen ed	Gen ed	MB	Combination	Team
	2	7	Female	At risk	Descript	N/A	Gen ed	MB	Combination	Team
	3	7	Male	At risk	Descript	N/A	Gen ed	MB	Combination	Team
	4	7	Female	LD	Descript	N/A	Gen ed	MB	Combination	Team
Filter & Horner (2009)	1	9	Male	LD	FA	Combination	Gen ed	CR	Researcher	Teacher
	2	9	Male	At risk	FA	Gen ed	Gen ed	CR	Researcher	Teacher
Grandy & Peck (1997)	1	6	Male	At risk	FA	Gen ed	Gen ed	MB	Prac student	Prac student
Hagan-Burke, et al. (2007)[a]	1	8	Male	At risk	FA	Gen ed	Gen ed	AB	Combination	Teacher

TABLE 2
Continued

Articles	Student	Age	Gender	Disability	FA	Assessment Setting	Intervention Setting	Experimental Design	Assessment Agent	Intervention Agent
Hoff et al. (2005)	1	12	Male	ADHD	FA	Gen ed	Gen ed	ABAB	Teacher	Teacher
Ingram et al. (2005)	1	11	Male	At risk	Descript	Gen ed	Gen ed	CR	Researcher	Teacher
	2	11	Male	At risk	Descript	Gen ed	Gen ed	CR	Researcher	Teacher
Jolivette et al. (2001)	1	7	Male	EBD	Descript	N/A	SPED	MB	Researcher	Teacher
	2	7	Male	EBD	Descript	SPED	SPED	MB	Researcher	Teacher
	3	7	Male	EBD	Descript	SPED	SPED	ABAB	Researcher	Teacher
Jones et al. (2000)	1	8	Male	ADHD	FA	Lab school	Lab school	ABAB	Researcher	Researcher
Kamps et al. (2006)	1	7	Female	At risk	FA	Gen ed	Gen ed	ABAB	Teacher	Teacher
	2	7	Male	At risk	FA	Gen ed	Gen ed	ABAB	Teacher	Teacher
Kennedy et al. (2001)	1	8	Male	EBD	Descript	N/A	Gen ed	MB	Researcher	Teacher
	2	6	Male	ADHD	Descript	N/A	Gen ed	MB	Researcher	Teacher
	3	6	Female	At risk	Descript	N/A	Gen ed	MB	Researcher	Teacher
Kern et al. (1994)	1	11	Male	EBD	FA	SPED	SPED	MB	Teacher	Teacher
Lane et al. (2006)	1	7	Male	At risk	Descript	N/A	Gen ed	ABAB	Teacher	Teacher
	2	7	Female	At risk	Descript	N/A	Gen ed	ABAB	Teacher	Teacher
Lane et al. (2007a)	1	7	Female	At risk	Descript	N/A	Gen ed	ABAB	Teacher	Teacher
	2	7	Male	At risk	Descript	N/A	Gen ed	CR	Teacher	Teacher
Lane et al. (2007b)	1	14	Male	LD	Descript	N/A	Gen ed	ABAB	Researcher	Teacher
Lane et al. (2007c)	1	7	Male	ADHD	Descript	SPED	SPED	ABAB	Team	Teacher
	2	6	Male	ADHD	Descript	N/A	Multiple settings	MB	Team	Team
Lane et al. (2007d)	1	6	Male	At risk	Descript	N/A	Multiple settings	MB	Combination	Teacher
Lane et al. (2009)	1	6	Male	At risk	Descript	N/A	Gen ed	ABAB	Researcher	Teacher
	2	6	Male	EBD	Descript	N/A	Gen ed	ABAB	Researcher	Teacher
Lewis & Sugai (1996)	1	6	Male	Language disorder	FA	Gen ed	Gen ed	CR	Researcher	Teacher
Liaupsin et al. (2006)	1	14	Female	At risk	Descript	N/A	Multiple settings	MB	Researcher	Teacher
Lo & Cartledge (2006)	1	7	Male	At risk	Descript	N/A	SPED	MB	Combination	Teacher
	2	8	Male	At risk	Descript	N/A	SPED	MB	Combination	Teacher
	3	9	Male	ADHD	Descript	N/A	SPED	MB	Combination	Teacher
	4	9	Male	EBD	Descript	N/A	SPED	MB	Combination	Teacher

TABLE 2
Continued

Articles	Student	Age	Gender	Disability	FA	Assessment Setting	Intervention Setting	Experimental Design	Assessment Agent	Intervention Agent
Maag & Larson (2004)	1	10	Male	EBD	Descript	N/A	SPED	MB	Teacher	Teacher
	2	10	Male	LD	Descript	N/A	SPED	MB	Teacher	Teacher
Magee & Ellis (2000)	1	7	Male	ADHD	FA	Isolated	Isolated	MB	Researcher	Researcher
March & Horner (2002)	1	13	Male	EBD	Descript	N/A	SPED	MB	Researcher	Teacher
	2	13		At risk	Descript	N/A	Gen ed	MB	Researcher	Teacher
	3	12	Female	LD	Descript	N/A	SPED	MB	Researcher	Teacher
McComas et al. (2002)	1	9	Male	LD	FA	SPED	SPED	AB	Researcher	Researcher
McLaren & Nelson (2009)	1	3	Male	At risk	Descript	N/A	Preschool	ABAB	Researcher	Teacher
	2	3	Male	At risk	Descript	N/A	Preschool	ABAB	Researcher	Teacher
	3	3	Male	At risk	Descript	N/A	Preschool	ABAB	Researcher	Teacher
Meyer (1999)	1	7	Male	EBD	FA	Isolated	Isolated	CR	Researcher	Researcher
	2	7	Male	EBD	FA	Isolated	Isolated	CR	Researcher	Researcher
	3	9	Female	EBD	FA	Isolated	Isolated	CR	Researcher	Researcher
	4	9	Male	EBD	FA	Isolated	Isolated	CR	Researcher	Researcher
Mueller et al. (2003)	1	7	Female	At risk	FA	SPED	SPED	ABAB	Teacher	Teacher
	2	7	Male	At risk	FA	Gen ed	Gen ed	ABAB	Teacher	Teacher
	3	7	Male	Hearing impaired	FA	SPED	SPED	ABAB	Teacher	Teacher
Packenham et al. (2004)	1	8	Female	At risk	Descript	N/A	Gen ed	MB	Teacher	Teacher
	2	9	Male	At risk	Descript	N/A	Gen ed	MB	Teacher	Teacher
Newcomer & Lewis (2004)	1	9	Male	OHI	FA	Gen ed	Gen ed	MB	Researcher	Teacher
	2	11	Male	At risk	FA	Gen ed	Gen ed	MB	Researcher	Teacher
	3	11	Female	At risk	FA	Gen ed	Gen ed	MB	Researcher	Teacher
Park & Scott (2009)[a]	1	5	Male	At risk	FA	Preschool	Preschool	ABAB	Researcher	Teacher
	2	4	Male	At risk	FA	Preschool	Preschool	ABAB	Researcher	Teacher
	3	4	Female	At risk	FA	Preschool	Preschool	ABAB	Researcher	Teacher
Patterson (2009)	1	15	Male	At risk	Descript	N/A	Gen ed	ABAB	Teacher	Teacher
Payne et al. (2007)	1	11	Female	LD	FA	SPED	SPED	CR	Researcher	Teacher
	2	10	Female	LD	FA	SPED	SPED	CR	Researcher	Teacher
	3	9	Male	EBD	FA	SPED	SPED	CR	Researcher	Teacher

TABLE 2
Continued

Articles	Student	Age	Gender	Disability	FA	Assessment Setting	Intervention Setting	Experimental Design	Assessment Agent	Intervention Agent
Penno et al. (2000)	1	13	Male	EBD	FA	SPED	SPED	AB	Combination	Teacher
	2	14	Male	EBD	FA	SPED	SPED	AB	Combination	Teacher
	3	14	Male	EBD	FA	SPED	SPED	AB	Combination	Teacher
Preciado et al. (2008)	1	7	Male	At risk	FA	Gen ed	Isolated	MB	Researcher	Grad student
	2	7	Male	LD	FA	Gen ed	Isolated	MB	Researcher	Grad student
	3	8	Female	At risk	FA	Gen ed	Isolated	MB	Researcher	Grad student
Shumate & Wills (2010)	1	7	Male	At risk	FA	Gen ed	Gen ed	MB	Teacher	Teacher
	2	8	Male	At risk	FA	Gen ed	Gen ed	MB	Teacher	Teacher
	3	7	Female	At risk	FA	Gen ed	Gen ed	MB	Teacher	Teacher
Skinner et al. (2009)	1	6	Male	ADHD	FA	Gen ed	Gen ed	AB	Teacher	Teacher
Smith & Sugai (2000)	1	13	Male	EBD	Descript	N/A	SPED	ABAB	Researcher	Teacher
Stage et al. (2006)	1	5	Male	ADHD	FA	Gen ed	Gen ed	AB	Researcher	Teacher
	2	6	Male	EBD	FA	SPED	SPED	AB	Researcher	Teacher
	3	14	Female	ADHD	FA	SPED	SPED	AB	Researcher	Teacher
Stahr et al. (2006)	1	9	Male	OHI	Descript	N/A	SPED	ABAB	Researcher	Teacher
Stichter et al. (2005) [a]	1	11	Male	EBD	FA	SPED	SPED	ABA	Researcher	Teacher
	2	11	Male	EBD	FA	SPED	SPED	ABA	Researcher	Teacher
Stichter et al. (2004) [a]	1	7	Male	EBD	FA	SPED	Gen ed	ABAB	Researcher	Teacher
Storey et al. (1994)	1	6	Male	At risk	Descript	N/A	Gen ed	ABAB	Researcher	Teacher
Trussell et al. (2008)	1	11	Male	EBD	FA	SPED	SPED	MB	Researcher	Teacher
	2	8	Male	EBD	FA	SPED	SPED	MB	Researcher	Teacher
	3	7	Male	EBD	FA	SPED	SPED	CR	Researcher	Teacher
Turton et al. (2007)	1	16	Female	EBD	Descript	N/A	Gen ed	ABAB	Researcher	Teacher
Umbreit (1995)	1	8	Male	ADHD	FA	Gen ed	Gen ed	MB	Combination	Teacher
Umbreit & Blair (1997)[a]	1	4	Male	EBD	FA	Preschool	Preschool	MB	Researcher	Teacher
Van Camp et al. (2000)	1	5	Male	At risk	FA	Isolated	Isolated	AB	Researcher	Researcher
Wood et al. (2007)	1	8	Male	At risk	Descript	Gen ed	Gen ed	CR	Researcher	Teacher

A total of 18 models were run separately using HLM 7.0 based on dichotomous comparisons of the categorical predictors described in *Table 4*, and those results were aggregated to a single table. None of the models with predictor variables significantly impacted the overall variance estimates, calculated using the model comparison χ^2 test within the HLM software. Therefore, only the random effects from the full model without predictors are reported. The results of the HLM meta-analysis are presented in *Table 5*. Because the outcome variable was percentage, interpretation of the results is in the percentage scale.

Full Model without Predictors

The mean estimate of students' percentage of problem behavior during baseline was 48.63% and the rate of change from baseline to intervention (mean shift) was −34.26%. At the session level, the rate of change by session during intervention was −0.52%, which was twice that of the session change during baseline (0.27), indicating a statistically significant downward slope in the intervention phase. The random effects within the full model in *Table 5* indicate that there was statistically significant variability between students in the initial status (mean baseline value), rate of change, and rate of change from baseline to intervention. This suggests that, although there was a statistically significant intervention effect, there was significant variability between students' intervention effects.

Moderating Variables

Across all fixed effects, none of the student-level characteristics were significant moderators of effect, indicating that the fixed effects were not systematically impacted by any student-level characteristics. At the assessment and intervention level, a number of moderators had a statistically significant impact on the fixed effects. Functional analysis was a statistically significant moderator, with a β coefficient of 7.93. The sign of the coefficient indicates that studies that utilized only descriptive assessments had a mean shift of 26.33% (−34.26 to 7.93), while studies that utilized experimental assessments had a mean shift of 34.26%. The difference was statistically significant. Interventions conducted in general education classrooms were also found to be

TABLE 2
Continued

Articles	Student	Age	Gender	Disability	FA	Assessment Setting	Intervention Setting	Experimental Design	Assessment Agent	Intervention Agent
Wright-Gallo et al. (2006)	1	14	Male	EBD	FA	SPED	SPED	ABAB	Researcher	Teacher
	2	12	Male	EBD	FA	SPED	SPED	ABAB	Researcher	Teacher

Note: Descript = descriptive assessment procedures; SPED = special education classroom; Gen ed = general education classroom; MB = multiple baseline; CR = complex reversal (reversal pattern with more than ABAB reversals); Para = paraprofessional; OHI is Other Health Impairments and FA is Functional Analysis or Descriptive Assessment.
[a]Studies utilizing structural analysis assessment procedures.

TABLE 3
Single-Subject Effect Size Results

Results	Student Level (n = 146)		Outcome Level (n = 206)	
	PND	SMD_{All}	PND	SMD_{All}
Mean	73.25	2.97	70.88	2.97[a]
Minimum	0	0.33	0	0.02
Maximum	100	19.24	100	19.24
Standard deviation	33.00	4.10	35.27	2.99
Pearson's *r*		.432**		.477**

Note: PND = percentage of nonoverlapping data points; SMD_{All} = standard mean difference all data. Pearson's *r* is the correlation between PND and SMD_{All}.
[a]The exact same values at the student and outcome level are accurate.
**$p < .01$.

significant predictors. During baseline, the dependent variable was 12.62% less in the general education classroom than in other intervention settings. The intervention effect was also impacted by the general education classroom setting, indicating that the mean shift was 13.11% less than in other intervention settings. Finally, intervention type was a significant predictor of rate of change during baseline and the rate of change across sessions between baseline and intervention. However, it should be noted that 90.5% of all the studies used positive procedure interventions. At the study level, the quality indicator score did not have an impact on any of the fixed effects, while the presence of significant autocorrelation was associated with a 12% increase in intervention effect. This result makes intuitive sense because significant autocorrelation signifies stable data results, not great variability. If an intervention effectively reduced the dependent variable, its effectiveness would be associated with a stable result, (i.e., a reduction to a consistent absence of an aberrant behavior).

It should be noted that other large coefficient values in both initial status and intervention effects reported in *Table 5* were not statistically significant. Those values were not statistically significant because of the standard error in their scores and because the distribution of probable values included 0.

Overall Effect

There was no significant change in the within- or between-variance results from the inclusion of the predictor variables. Therefore, an overall effect size was calculated using the fixed effects coefficients from the full model without predictors as those values represent overall effectiveness. Based on the mean baseline value of 48.63% and the intervention effect, −34.26%, FBA-based interventions for students with or at risk for EBD reduced problem behaviors by 70.5% (34.26/48.63).

Discussion

This study was designed to examine the empirical evidence of FBA-based interventions by conducting an HLM meta-analysis of FBA-based intervention research for students with or at risk for EBD. Unlike other meta-analytic SSD methods, HLM meta-analysis provides values and significance for the three key components of visual analysis—mean shift, trend, and variability (Parsonson & Baer, 1992)—which are elaborated below and document the omnibus impact of FBA-based interventions.

Overall FBA-Based Intervention Effects

Mean Shift

The mean shift, or rate of change from baseline to intervention, was −34.26%, interpreted as the percentage of change from initial status, or baseline (48.63%) to intervention. These findings indicate that FBA-based interventions reduce the percentage of the dependent variable by 70.5% from baseline to intervention. The mean shift effect was statistically significant. Overall, FBA-based interventions for students with or at risk for EBD, on average, reduced problem behaviors by 70.5%. This finding is in line with the findings of Goh and Bambara (in press) and Marquis et al. (2000), both of which calculated effects similar to mean shift. Comparison with the Gresham et al. (2004) study is

TABLE 4
Dichotomous Model Comparisons

Categorical Variables	Dummy Coding Scheme
Disability	1. ADHD compared to all other disabilities 2. EBD compared to all others disabilities 3. At risk compared to all others disabilities
Functional analysis	1. Functional analysis procedures compared to descriptive assessments
Assessment setting	1. Isolated room compared to other settings 2. School classrooms compared to other settings 3. Preschool compared to other settings 4. General education classrooms compared to other settings 5. Special education classrooms compared to other settings
Intervention setting	1. Isolated room compared to other settings 2. School classrooms compared to other settings 3. Preschool compared to other settings 4. General education classrooms compared to other settings 5. Special education classrooms compared to other settings
Assessment agent	1. Researcher compared to other agents 2. Teacher (including paraprofessionals) compared to other agents
Intervention agent	1. Researcher compared to other agents 2. Teacher (including paraprofessionals) compared to other agents
Behavior topography	1. Multiple behaviors compared to other behaviors

difficult because it included non-FBA intervention studies, but the magnitude of effect from the FBA-based studies was very similar.

Trend

The mean rate of session change during the baseline phase was 0.27, indicating that the baseline values had a positive, or increasing, slope. The β coefficient slope value for baseline was not statistically significant, indicating that the real slope value could be 0. The mean rate of session change from baseline to intervention was −0.52, indicating a negative, or decreasing, slope. The intervention slope was statistically significant. These results indicate that the baseline slope was flat while the intervention slope had a significant declining trend.

Variability

The variance components within the model indicated that there was significant variability within and between students. Both the baseline phase and intervention phase rate of change were significant, meaning there was significant variability in the data pattern, or spread of data points, in both baseline and intervention phases. The between-student variability suggests that although the intervention was effective, the results varied between students.

Based on the final full model without predictors, the mean shift may be the best predictor of overall effectiveness and, in combination with the negative significant slope for intervention, support the hypothesis that FBA-based interventions are effective. These results, in combination with the two effect size results, empirically support FBA-based interventions as an effective practice.

Student Characteristics

The age range of students was large (3–16 years), with a mean of 7.92, suggesting that the majority of studies were conducted with elementary-age students. Within the models, the age of the student did not significantly impact the mean shift from baseline to intervention or the rate of change. Based on these results, FBA-based interventions were effective across the age range of included students. Gender was also modeled, but only a small number of females (n = 25) were included in the sample. Overall, gender was not a significant predictor of mean shift or trend.

Although not significant, a difference in the coefficient size was present between disability categories. These differences are discussed within the context of a visual analysis approach where slight differences in data patterns are examined for practical significance, not statistical significance (Hopkins, Cole, & Mason, 1998). The β coefficients

TABLE 5
HLM-Meta-Analysis of FBA-Based Interventions for Student with or At Risk for Emotional and/or Behavioral Disorders

Fixed Effects	Initial Status (π_0)	Baseline Slope (π_1)	Intervention Effect (π_2)	Intervention Slope (π_3)
Full model w/o predictors				
Intercept	48.63***	0.27	−34.26***	−0.52***
Moderating variables				
Age	−0.53	0.04	0.80	−0.01
Gender	−3.46	0.56	3.13	−0.55
ADHD	7.97	0.35	−5.67	−0.57
EBD	9.31	0.61	−6.48	−0.73
At risk	0.41	0.45	1.51	−0.44
Behavior topography	0.53	0.23	−0.55	−0.14
Functional analysis	−4.01	0.04	7.93*	−0.01
Isolated assessment	−15.01	−0.63	13.97	0.18
Isolated intervention	12.35	0.31	−9.59	−0.25
Gen ed assessment	3.04	0.63	−5.45	−0.79
Gen ed intervention	−12.62*	−0.23	13.11**	0.47
SPED assessment	−3.39	0.24	4.75	−0.43
SPED intervention	−9.52	0.18	8.45	0.05
Researcher assessment	7.48	−0.03	−5.78	−0.02
Researcher intervention	−3.55	−0.01	6.69	−0.09
Teacher assessment	7.58	−0.50	−7.51	0.56
Teacher intervention	−5.77	−0.01	11.95	−0.06
Intervention type	−2.08	1.23**	−0.68	−1.24*
Quality indicators	0.14	−0.05	0.05	0.10
Autocorrelation	10.49	0.00	−11.65**	0.01
Random effects				
Variance	431.971***	0.238***	410.862***	7.141***
Residual variance	234.136***			

Note: Gen ed = general education classroom; SPED = special education classroom.
*$p < .05$; **$p < .01$; ***$p < .001$.

for ADHD, EBD, and nonidentified hint at a slight difference in mean shift between the disability categories. Hierarchically, FBA-based interventions may be slightly more effective for students with EBD, then ADHD, and then at-risk students based on the data. However, regardless of disability, the significance of mean shift remains.

Assessment and Intervention Characteristics

As indicated above, there are divergent opinions within the FBA literature about the procedures and processes for assessment and intervention development (Gresham et al., 2001; Sasso et al., 2001). To explore *some* of the procedural questions about the effectiveness of FBA-based intervention, assessment and intervention characteristics were included as predictor variables among the fixed effects.

Descriptive assessment significantly predicted difference in mean shift, with a β coefficient of 7.93. In other words, FBA-based interventions that do not use functional analysis in the development of an intervention appear to be less effective at reducing problem behaviors. Scott and Kamps (2007) highlight the contentious state of current knowledge about whether the confirmation of the hypothesized function or antecedent of a behavior is necessary for FBA-based intervention development

but suggest that descriptive methods may be the most "efficient and sufficient in many cases" (p. 152). The results of this study indicate that FBA-based interventions are significantly effective regardless of assessment procedure; however, functional analysis may be more effective for students with or at risk for EBD in schools. A number of issues could account for the differences found between the assessment procedures, including more accurate selection of interventions when using functional analysis and the fact that experimental assessment studies were more carefully designed and executed. The results suggest the need for further research.

The findings from the models examining assessment and intervention procedures conducted in general education and special education classrooms indicate that the mean shift reduction between baseline and intervention was statistically significantly different for interventions conducted in general education classrooms. The mean shift β coefficient for interventions conducted in general education classrooms was 13.11. This result is not surprising because, in theory, to remain in a general education classroom, students would not exhibit the same magnitude of behavior as those students in self-contained settings. Therefore, the percentage reduction is possibly impacted by the reduced percentage of the target behavior within a general education classroom.

Models were constructed to examine whether differences could be found between researchers and teachers in the assessment and intervention process. The findings indicated that there were no differences between either of the agents among all the fixed effects. An issue within the FBA literature is whether teachers can effectively conduct FBAs in natural settings (Drasgow, Yell, Bradley, & Shriner, 1999; Johnston & O'Neill, 2001; Scott & Kamps, 2007). The findings of this study suggest that there are no statistically significant differences in FBA-based results between researchers and teachers in the literature. However, this finding must be couched in the sample size of teacher-implemented assessments, which was only 19% of the sample. Although the proportions within the dummy variables do not bias the results, there is a need for more studies to further examine this issue. Nonetheless, the findings provide support for teachers' ability to conduct FBAs in schools.

The last assessment and intervention characteristics modeled examined whether behavior topography and type of intervention predicted differences in the fixed effects. No significant moderating effects were found for behavior topography. For type of intervention, positive procedures far outnumbered the other intervention types (punishment, extinction, and combination) used in the literature (90.5% compared with 9.5%). Yet a significant difference was found in trend between positive procedures and the extinction–punishment interventions. The β coefficient, or slope, for extinction–punishment was 0.57, meaning that it was increasing each session. For positive procedures, the β coefficient, or slope, was −1.24, meaning that it decreased the percentage of a behavior each session.

Study-Level Characteristics

The quality indicator score was modeled to examine whether differences across fixed effects were systematically impacted by study quality, finding there was no impact. This suggests that the results were not impacted by study quality, keeping in mind that most of the studies received moderately high quality scores. The results related to autocorrelation indicate that significant autocorrelation resulted in a more effective intervention mean shift value. In other words, the more lag-1 autocorrelation, the greater the mean shift, which is expected because more lag-1 autocorrelation signified less variability in the dependent variable; less variability means fewer outlier values driving the mean upward.

Limitations

Although many checks were included to ensure the validity of the findings, a number of limitations necessitate highlighting. The study inclusion criteria were designed to be as inclusive as possible but required a number of specific study characteristics for the analysis. The replicable baseline criteria excluded some studies that did not develop an intervention or discuss the development of an intervention. Additionally, the inclusion of studies that compared FBA interventions with other interventions might find different effect sizes and broaden our knowledge about FBA interventions versus other types of interventions (e.g., Check-In/Check-Out). Further, as with any meta-analysis, the "file drawer" effect, whereby studies absent of effect are not published and not included, is a concern.

Our focus was to identify the state of the peer-reviewed literature, but broader inclusion criteria, including "gray" literature, could have increased the sample size and impacted results.

A number of limitations were also present within the analysis. The first issue is the reliability and validity of the quality indicator score; using an instrument with valid psychometric properties would increase validity of the score. An overall limitation of HLM meta-analysis is the need for types of SSD to be similar because of the aggregation approach. In this study, only studies with baselines and interventions were used. Further research validating the HLM meta-analysis procedure and how results compare with other SSD effect sizes is also necessary. A more complete examination of the small correlation coefficient between the effect sizes should also be a target for future research.

Implications for Practice

This study provides two primary implications for practice. First, the results suggest that FBA-based interventions in schools for students with EBD can be implemented effectively. However, the number of studies absent of researcher influence in the assessment and intervention process was very limited. The implication is that educators can conduct these procedures but that they may need more support and training at the pre- and in-service levels (Scott & Kamps, 2007). These assessments and interventions are being conducted in general education settings; therefore, general educators should have the requisite knowledge to support the process or conduct stand-alone FBA assessments. The findings indicate that FBA-based interventions are very effective; therefore, in-service professionals should have the requisite skills and knowledge to conduct the assessment and develop the interventions in order to help students with behavioral concerns.

The second implication for practice was the finding of hierarchy in assessment procedures. Although FBA-based interventions as a whole were found to be effective, descriptive assessments alone were less effective. For students with the most severe behavior, the most effective assessment-to-intervention processes should be implemented, and the results suggest functional analysis. Pre- and in-service professionals need comprehensive training on the implementation of experimental manipulations in natural settings. Although further research is needed to examine effective training models, the results of this study indicate that school professionals should have the skills to conduct experimental manipulations in schools.

Conclusion

This study reported on a novel meta-analysis procedure for SSD and identified the effectiveness of FBA-based interventions for students with EBD. The findings replicate those of Goh and Bambara (in press) and support contentions that FBA-based interventions are an evidence-based practice. Although further research is warranted, the results reported here supply the field with evidence supporting the effective application of FBA-based interventions for students with EBD in schools.[1]

NOTE

1. Please contact the first author for a list of the included studies.

REFERENCES

Bessette, K. K., & Wills, H. P. (2007). An example of an elementary school paraprofessional-implemented functional analysis and intervention. *Behavioral Disorders, 32*(3), 192–210.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis.* West Sussex, UK: John Wiley & Sons.

Busk, P. L., & Serin, S. R. (1992). Meta-analysis for single-case research. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case research design and analysis: New directions for psychology and education* (pp. 187–212). Hillsdale, NJ: Erlbaum Associates.

Campbell, J. M. (2003). Efficacy of behavioral interventions for reducing problem behavior in persons with autism: A quantitative synthesis of single-subject research. *Research in Developmental Disabilities, 24*(1), 120–138, doi:10.1016/S0891-4222(03)00014-3

Carr, E. G., Horner, R. H., Turnbull, A. P., Marquis, J. G., McLaughlin, D. M., McAtee, M. L., & ...Braddock, D. (1999). *Positive behavior support with people with developmental disabilities: A research synthesis.* Washington, D. C.: American Association on Mental Retardation.

Christensen, L., Young, K. R., & Marchant, M. (2007). Behavioral intervention planning: Increasing appropriate behavior of a socially withdrawn student. *Education and Treatment of Children, 30*(4), 81–103, doi:10.1353/etc.2007.0022

Conroy, M. A., & Stichter, J. P. (2003). The application of antecedents in the functional assessment process: Existing research, issues, and recommendations. *The Journal of Special Education, 37*(1), 15–25, doi:10.1177/00224669030370010201

Drasgow, E., Yell, M. L., Bradley, R., & Shriner, J. G. (1999). The IDEA amendments of 1997: A school-wide model for conducting functional behavioral assessments and developing behavior intervention plans. *Education and Treatment of Children, 22*(3), 244–266, Retrieved from http://www.educationandtreatmentofchildren.net/

Dunlap, G., Kern-Dunlap, L., Clarke, S., & Robbins, F. R. (1991). Functional assessment, curricular revision, and severe behavior problems. *Journal of Applied Behavior Analysis, 24*(2), 387–397, doi:10.1901/jaba.1991.24–387

Edwards, W. H., Magee, S. K., & Ellis, J. (2002). Identifying the effects of idiosyncratic variables on functional analysis outcomes: A case study. *Education and Treatment of Children, 25*(3), 317–330, Retrieved from http://www.educationandtreatmentofchildren.net/

Ellis, J., & Magee, S. K. (1999). Determination of environmental correlates of disruptive classroom behavior: Integration of functional analysis into public school assessment process. *Education and Treatment of Children, 22*(3), 291–316, Retrieved from http://www.educationandtreatmentofchildren.net/

Faith, M. S., Allison, D. B., & Gorman, B. S. (1996). Meta-analysis of single-case research. In R. D. Franklin, D. B. Allison, & B. S. Gorman (Eds.), *Design and analysis of single-case research* (pp. 245–277). Hillsdale, NJ: Lawrence Erlbaum.

Fox, J. J., & Davis, C. A. (2005). Functional behavior assessment in schools: Current research findings and future directions. *Journal of Behavioral Education, 14*(1), 1–4, doi:10.1007/s10864-005-0957-0

Goh, A. E., & Bambara, L. M. ((in press)). Individualized positive behavior support in school settings: A meta-analysis. *Remedial and Special Education, XX*(x), 1–16, doi:10.1177/0741932510383990

Grandy, S. E., & Peck, S. M. (1997). The use of functional assessment and self-management with a first grader. *Child & Family Behavior Therapy, 19*(1), 29–43, doi:10.1300/J019v19n02_03

Gresham, F. M., McIntyre, L. L., Olson-Tinker, H., Dolstra, L., McLaughlin, V., & Van, M. (2004). Relevance of functional behavioral assessment research for school-based intervention and positive behavioral support. *Research in Developmental Disabilities, 25*(1), 19–37, doi:10.1016/j.ridd.2003.04.003

Gresham, F. M., Watson, T. S., & Skinner, C. H. (2001). Functional behavioral assessment: Principles, procedures, and future directions. *School Psychology Review, 30*(2), 156–172, Retrieved from http://www.nasponline.org/publications/spr/about.aspx

Guo, S., & Hussey, D. (1999). Analyzing longitudinal data: A three-level hierarchical linear model. *Social Work Research, 23*(4), 258–268, doi:10.1093/swr/23.4.258

Hoff, K. E., Ervin, R. A., & Friman, P. C. (2005). Refining functional behavioral assessment: Analyzing the separate and combined effects of hypothesized controlling variables during ongoing classroom routines. *School Psychology Review, 34*(1), 45–57, Retrieved from http://www.nasponline.org/publications/spr/about.aspx

Hopkins, B. L., Cole, B. L., & Mason, T. L. (1998). A critique of the usefulness of inferential statistics in applied behavior analysis. *The Behavior Analyst, 21*(1), 125–137, Retrieved from http://www.abainternational.org/TBA.asp

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Worley, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children, 71*(2), 165–179, Retrieved from http://www.cec.sped.org/Content/NavigationMenu/Publications2/ExceptionalChildren/default.htm

Hudson, S., Lewis, T. J., Stichter, J. P., & Johnson, N. W. (2011). Putting quality indicators to the test: An examination of 30 years of research. *Journal of Emotional and Behavioral Disorders, 19*(3), 143–155, doi:10.1177/1063426610364986

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1982/1994). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis, 27*(2), 3–20, doi:10.1016/0270-4684(82)90003-9

Jenson, W. R., Clark, E., Kircher, J. C., & Kristjansson, S. D. (2007). Statistical reform: Evidence-based practice, meta-analyses, and single subject designs. *Psychology in the Schools, 44*(5), 483–493, doi:10.1002/pits.20240

Johnston, S. S., & O'Neill, R. E. (2001). Searching for effectiveness and efficiency in conducting functional assessments: A review and proposed process for teachers and other practitioners. *Focus on Autism and Other Developmental Disabilities, 16*(4), 205–214, doi:10.1177/108835760101600402

Kamps, D. M., Wendland, M., & Culpepper, M. (2006). Active teacher participation in functional behavior assessment for students with emotional and behavioral disorders risks in general education classrooms. *Behavioral Disorders, 31*(2), 128–146.

Kavale, K. A. (1984). Potential advantages of the meta-analysis technique for research in special education. *Journal of Special Education, 18*(1), 61–72, doi:10.1177/002246698401800107

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

Lane, K. L., Kalberg, J. R., & Shepcaro, J. C. (2009). An examination of the evidence base for function-based interventions for students with emotional or behavioral disorders attending middle and high schools. *Exceptional Children, 75*(3), 321–340.

Lane, K. L., Rogers, L. A., Parks, R. J., Weisenbach, J. L., Mau, A. C., & Merwin, M. T., et al. (2007). Function-based interventions for students who are nonresponsive to primary and secondary prevention efforts: Illustrations at the elementary and middle school levels. *Journal of Emotional and Behavioral Disorders, 15*(3), 169–183.

Lane, K. L., Umbreit, J., & Beebe-Frankenberger, M. E. (1999). Functional assessment research on students with or at risk for EBD: 1990 to the present. *Journal of Positive behavior Interventions, 1*(1), 101–111, doi:10.1177/109830079900100204

Liaupsin, C., Umbreit, J., Ferro, J. B., Urso, A., & Upreti, G. (2006). Improving academic engagement through systematic, function-based intervention. *Education and Treatment of Children, 29*(4), 573–591, Retrieved from http://www.educationandtreatmentofchildren.net/

Maag, J. W., & Larson, P. J. (2004). Training a general education teacher to apply functional assessment. *Education and Treatment of Children, 27*(1), 26–36, Retrieved from http://www.educationandtreatmentofchildren.net/

Maggin, D. M., Swaminathan, H., Rogers, H. J., O'Keefe, B. V., Sugai, G., & Horner, R. H. (2011). A generalized least squares regression approach for computing effect sizes in single-case research: Application examples. *Journal of School Psychology, 49,* 301–321, doi:10.1016/j.jsp.2011.03.004

Marquis, J. G., Horner, R. H., Carr, E. G., Turnbull, A. P., Thompson, M., & Behrens, G. A., et al. (2000). A meta-analysis of positive behavior support. In R. Gersten, E. P. Schiller, & B. J. Vaughn (Eds.), *Contemporary special education research: Syntheses of the knowledge base on critical instructional issues* (pp. 137–178). Mahwah, NJ: Lawrence Erlbaum Associates.

Moore, J. W., Edwards, R. P., Sterling-Turner, H. E., Riley, J., DuBard, M., & McGeorge, A. (2002). Teacher acquisition of functional analysis methodology. *Journal of Applied Behavior Analysis, 35*(1), 73–77, doi:10.1901/jaba.2002.35-73

Nagler, E. M., Rindskopf, D. M., & Shadish, W. R. (2008). Analyzing data from small *N* designs using multilevel models: A procedural handbook. Washington, D. C.: U.S. Department of Education.

Neilsen, S. L., & Mcevoy, M. A. (2004). Functional behavioral assessment in early education settings. *Journal of Early Intervention, 16*(2), 115–131, doi:10.1177/105381510402600204

Nelson, J. R., Roberts, M. L., Mathur, S. R., & Rutherford, R. B. (1999). Has public policy exceeded our knowledge base? A review of the functional behavioral assessment literature. *Behavioral Disorders, 24*(2), 169–179.

Park, K. L., & Scott, T. M. (2009). Antecedent-based interventions for young children at risk for emotional and behavioral disorders. *Behavioral Disorders, 34*(4), 196–211.

Parsonson, B. S., & Baer, D. M. (1992). The visual analysis of data, and current research into the stimuli controlling it. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case research design and analysis: New directions for psychology and education* (pp. 15–40). Hillsdale, NJ: Lawrence Erlbaum Associates.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis procedures* (2nd ed.). Thousand Oaks, CA: Sage.

Raudenbush, S. W., Bryk, A. S., & Congdon, R. (2010). HLM 7 for Windows [Computer software]. Lincolnwood, IL: Scientific Software International, Inc.

Sasso, G. M., Conroy, M. A., Stichter, J. P., & Fox, J. J. (2001). Slowing down the bandwagon: The misapplication of functional assessment for students with emotional or behavioral disorders. *Behavioral Disorders, 26*(4), 282–296.

Scott, T. M., & Kamps, D. M. (2007). The future of functional behavioral assessment in school settings. *Behavioral Disorders, 32*(2), 146–157.

Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single-subject research: Methodology and validation. *Remedial and Special Education, 8*(2), 24–33, doi:10.1177/074193258700800206

Shadish, W. R., Brasil, I. C. C., Illingworth, D. A., Gilindo, R., Nagler, E. M., & Rindskopf, D. M. (2009). Using ungraph to extract data dfrom image files: Verifification of reliability and validity. *Behavior Research Methods, 41*(1), 177–183, doi:10.3758/BRM.41.1.177

Skinner, J. N., Veerkamp, M. B., Kamps, D. M., & Andra, P. R. (2009). Teacher and peer participation in functional analysis and intervention for a first grade student with attention deficit hyperactivity disorder. *Education and Treatment of Children, 32*(2), 243–266, Retrieved from http://www.educationandtreatmentofchildren.net/

Solnick, M. D., & Ardoin, S. P. (2010). A quantitative review of functional analysis procedures in public school settings. *Education and Treatment of Children, 33*(1), 153–175, Retrieved from http://www.educationandtreatmentofchildren.net/

Stahr, B., Cushing, D., Lane, K. L., & Fox, J. J. (2006). Efficacy of a functional-based intervention in decreasing off-task behavior exhibited by a student with ADHD. *Journal of Positive behavior Interventions, 8*(4), 201–211, doi:10.1177/10983007060080040301

Trussell, R. P., Lewis, T. J., & Stichter, J. P. (2008). The impact of targeted classroom interventions and function-based behavior interventions on problem behaviors of students with emotional/behavioral disorders. *Behavioral Disorders, 33*(3), 153–166.

Umbreit, J. (1995). Functional assessment and intervention in a regular classroom setting for the disruptive behavior of a student with attention deficit hyperactivity disorder. *Behavioral Disorders, 20*(4), 267–278.

Umbreit, J., & Blair, K. S. (1997). Using structural analysis to facilitate treatment of aggression and

noncompliance in a young child at risk for behavioral disorders. *Behavioral Disorders, 22*(1), 75–86.

Van den Noorgate, W., & Onghena, P. (2003). Combining single-case experimental data using hierarchical linear models. *School Psychology Quarterly, 18*(3), 325–346, doi:10.1521/scpq.18.3.325.22577

Van den Noortgate, W., & Onghena, P. (2007). The aggregation of single-case results using hierarchical linear models. *The Behavior Analyst, 8*(2), 196–208, Retrieved from http://www.abainternational.org/TBA.asp

Van den Noorgate, W., & Onghena, P. (2008). A multilevel meta-analysis of single-subject experimental design studies. *Evidence-based communication Assessment & Intervention, 2*(3), 142–151, doi:10.1080/17489530802505362

Wallace, M. D., Doney, J. K., Mintz-Resudek, C. M., & Tarbox, R. S. F. (2004). Training educators to implement functional analyses. *Journal of Applied Behavior Analysis, 37*(1), 89–92, doi:10.1901/jaba.2004.37-89

Watson, T. S., Ray, K. P., Turner, H. S., & Logan, P. (1999). Teacher-implemented functional analysis and treatment: A method for linking assessment to intervention. *School Psychology Review, 28*, 292–302, Retrieved from http://www.nasponline.org/publications/spr/about.aspx

Weber, K. P., Killu, K., Derby, K. M., & Barretto, A. (2005). The status of functional behavioral assessment (FBA): Adherence to standard practice in FBA methodology. *Psychology in the Schools, 42*(7), 737–744, doi:10.1002/pits.20108

Wolery, M., Busick, M., Reichow, B., & Barton, E. E. (2010). Comparison of overlap methods for quantitatively synthesizing single-subject data. *Journal of Special Education, 44*, 18–28, doi:10.1177/0022466908328009

Wood, B. K., Blair, K. S., & Ferro, J. B. (2009). Young children with challenging behavior: Function-based assessment and intervention. *Topics in Early Childhood Special Education, 29*(2), 68–78, doi:10.1177/0271121409337951

Yell, M. L. (2006). *The law and special education* (2nd ed.). Upper Saddle River, NJ: Pearson Merrill Prentice Hall.

AUTHORS' NOTE

Address correspondence to Nicholas A. Gage, Neag School of Education–CBER, 249 Glenbrook Road, Storrs, CT 06269-2064, USA. E-mail: Nicholas.gage@uconn.edu.

MANUSCRIPT

Initial Acceptance: 7/5/11
Final Acceptance: 12/29/11

APPENDIX A
Taxonomy of Hierarchical Linear Models

Taxonomy of Hierarchical Linear Models

Model	Level-1 Model	Level-2 Model	Mixed Model
A	$DV_{ti} = \pi_{0i} + e_{ti}$	$\pi_{0i} = \beta_{00} + r_0$	$DV_{ti} = \beta_{00} + r_{0i} + e_{ti}$
B	$DV_{ti} = \pi_{0i} + \pi_{1i}*(SESS1_{ti}) + e_{ti}$	$\pi_{0i} = \beta_{00} + r_{0i}$ $\pi_{1i} = \beta_{10} + r_{1i}$	$DV_{ti} = \beta_{00} + \beta_{10}*SESS1_{ti} + r_{0i} + r_{1i}*SESS1_{ti} + e_{ti}$
C	$DV_{ti} = \pi_{0i} + \pi_{1i}*(SESS1_{ti}) + \pi_{2i}*(PHASE_{ti}) + \pi_{3i}*(SXP_{ti}) + \pi_{4i}*(PXO_{ti}) + e_{ti}$	$\pi_{0i} = \beta_{00} + r_{0i}$ $\pi_{1i} = \beta_{10} + r_{1i}$ $\pi_{2i} = \beta_{20} + r_{2i}$ $\pi_{3i} = \beta_{30} + r_{3i}$ $\pi_{4i} = \beta_{40} + r_{4i}$	$DV_{ti} = \beta_{00} + \beta_{10}*SESS1_{ti} + \beta_{20}*PHASE_{ti} + \beta_{30}*SXP_{ti} + \beta_{40}*PXO_{ti} + r_{0i} + r_{1i}*SESS1_{ti} + r_{2i}*PHASE_{ti} + r_{3i}*SXP_{ti} + e_{ti}$
D	$DV_{ti} = \pi_{0i} + \pi_{1i}*(SESS1_{ti}) + \pi_{2i}*(PHASE_{ti}) + \pi_{3i}*(ORDER_{ti}) + \pi_{4i}*(SXPXO_{ti}) + e_{ti}$	$\pi_{0i} = \beta_{00} + r_{0i}$ $\pi_{1i} = \beta_{10} + r_{1i}$ $\pi_{2i} = \beta_{20} + r_{2i}$ $\pi_{3i} = \beta_{30} + r_{3i}$ $\pi_{4i} = \beta_{40} + r_{4i}$	$DV_{ti} = \beta_{00} + \beta_{10}*SESS1_{ti} + \beta_{20}*PHASE_{ti} + \beta_{30}*ORDER_{ti} + \beta_{40}*SXPXO_{ti} + r_{0i} + r_{1i}*SESS1_{ti} + r_{2i}*PHASE_{ti} + e_{ti}$
E	$DV_{ti} = \pi_{0i} + \pi_{1i}*(SESS1_{ti}) + \pi_{2i}*(PHASE_{ti}) + \pi_{3i}*(SXP_{ti}) + e_{ti}$	$\pi_{0i} = \beta_{00} + \beta_{01}*(AGE_i) + \beta_{02}*(GENDER1_i) + r_{0i}$ $\pi_{1i} = \beta_{10} + \beta_{11}*(AGE_i) + \beta_{12}*(GENDER1_i) + r_{1i}$ $\pi_{2i} = \beta_{20} + \beta_{21}*(AGE_i) + \beta_{22}*(GENDER1_i) + r_{2i}$ $\pi_{3i} = \beta_{30} + \beta_{31}*(AGE_i) + \beta_{32}*(GENDER1_i) + r_{3i}$	$DV_{ti} = \beta_{00} + \beta_{01}*AGE_i + \beta_{02}*GENDER1_i + \beta_{10}*SESS1_{ti} + \beta_{11}*AGE_i*SESS1_{ti} + \beta_{12}*GENDER1_i*SESS1_{ti} + \beta_{20}*PHASE_{ti} + \beta_{21}*AGE_i*PHASE_{ti} + \beta_{22}*GENDER1_i*PHASE_{ti} + \beta_{30}*SXP_{ti} + \beta_{31}*AGE_i*SXP_{ti} + \beta_{32}*GENDER1_i*SXP_{ti} + r_{0i} + r_{1i}*SESS1_{ti} + r_{2i}*PHASE_{ti} + r_{3i}*SXP_{ti} + e_{ti}$
F	$DV_{ti} = \pi_{0i} + \pi_{1i}*(SESS1_{ti}) + \pi_{2i}*(PHASE_{ti}) + \pi_{3i}*(SXP_{ti}) + e_{ti}$	$\pi_{0i} = \beta_{00} + \beta_{01}*(ADHD_i) + \beta_{02}*(EBD_i) + \beta_{03}*(NONIDENT_i) + r_{0i}$ $\pi_{1i} = \beta_{10} + \beta_{11}*(ADHD_i) + \beta_{12}*(EBD_i) + \beta_{13}*(NONIDENT_i) + r_{1i}$ $\pi_{2i} = \beta_{20} + \beta_{21}*(ADHD_i) + \beta_{22}*(EBD_i) + \beta_{23}*(NONIDENT_i) + r_{2i}$ $\pi_{3i} = \beta_{30} + \beta_{31}*(ADHD_i) + \beta_{32}*(EBD_i) + \beta_{33}*(NONIDENT_i) + r_{3i}$	$DV_{ti} = \beta_{00} + \beta_{01}*ADHD_i + \beta_{02}*EBD_i + \beta_{03}*NONIDENT_i + \beta_{10}*SESS1_{ti} + \beta_{11}*ADHD_i*SESS1_{ti} + \beta_{12}*EBD_i*SESS1_{ti} + \beta_{13}*NONIDENT_i*SESS1_{ti} + \beta_{20}*PHASE_{ti} + \beta_{21}*ADHD_i*PHASE_{ti} + \beta_{22}*EBD_i*PHASE_{ti} + \beta_{23}*NONIDENT_i*PHASE_{ti} + \beta_{30}*SXP_{ti} + \beta_{31}*ADHD_i*SXP_{ti} + \beta_{32}*EBD_i*SXP_{ti} + \beta_{33}*NONIDENT_i*SXP_{ti}$
G	$DV_{ti} = \pi_{0i} + \pi_{1i}*(SESS1_{ti}) + \pi_{2i}*(PHASE_{ti}) + \pi_{3i}*(SXP_{ti}) + e_{ti}$	$\pi_{0i} = \beta_{00} + \beta_{01}*(BEH\_TOP_i) + \beta_{02}*(COMORBID_i) + r_{0i}$ $\pi_{1i} = \beta_{10} + \beta_{11}*(BEH\_TOP_i) + \beta_{12}*(COMORBID_i) + r_{1i}$ $\pi_{2i} = \beta_{20} + \beta_{21}*(BEH\_TOP_i) + \beta_{22}*(COMORBID_i) + r_{2i}$ $\pi_{3i} = \beta_{30} + \beta_{31}*(BEH\_TOP_i) + \beta_{32}*(COMORBID_i) + r_{3i}$	$DV_{ti} = \beta_{00} + \beta_{01}*BEH\_TOP_i + \beta_{02}*COMORBID_i + \beta_{10}*SESS1_{ti} + \beta_{11}*BEH\_TOP_i*SESS1_{ti} + \beta_{12}*COMORBID_i*SESS1_{ti} + \beta_{20}*PHASE_{ti} + \beta_{21}*BEH\_TOP_i*PHASE_{ti} + \beta_{22}*COMORBID_i*PHASE_{ti} + \beta_{30}*SXP_{ti} + \beta_{31}*BEH\_TOP_i*SXP_{ti} + \beta_{32}*COMORBID_i*SXP_{ti} + r_{0i} + r_{1i}*SESS1_{ti} + r_{2i}*PHASE_{ti} + r_{3i}*SXP_{ti} + e_{ti}$
H	$DV_{ti} = \pi_{0i} + \pi_{1i}*(SESS1_{ti}) + \pi_{2i}*(PHASE_{ti}) + \pi_{3i}*(SXP_{ti}) + e_{ti}$	$\pi_{0i} = \beta_{00} + \beta_{01}*(FA_i) + r_{0i}$ $\pi_{1i} = \beta_{10} + \beta_{11}*(FA_i) + r_{1i}$ $\pi_{2i} = \beta_{20} + \beta_{21}*(FA_i) + r_{2i}$ $\pi_{3i} = \beta_{30} + \beta_{31}*(FA_i) + r_{3i}$	$DV_{ti} = \beta_{00} + \beta_{01}*FA_i + \beta_{10}*SESS1_{ti} + \beta_{11}*FA_i*SESS1_{ti} + \beta_{20}*PHASE_{ti} + \beta_{21}*FA_i*PHASE_{ti} + \beta_{30}*SXP_{ti} + \beta_{31}*FA_i*SXP_{ti} + r_{0i} + r_{1i}*SESS1_{ti} + r_{2i}*PHASE_{ti} + r_{3i}*SXP_{ti} + e_{ti}$
I	$DV_{ti} = \pi_{0i} + \pi_{1i}*(SESS1_{ti}) + \pi_{2i}*(PHASE_{ti}) + \pi_{3i}*(SXP_{ti}) + e_{ti}$	$\pi_{0i} = \beta_{00} + \beta_{01}*(NO\_MANIP_i) + r_{0i}$ $\pi_{1i} = \beta_{10} + \beta_{11}*(NO\_MANIP_i) + r_{1i}$ $\pi_{2i} = \beta_{20} + \beta_{21}*(NO\_MANIP_i) + r_{2i}$ $\pi_{3i} = \beta_{30} + \beta_{31}*(NO\_MANIP_i) + r_{3i}$	$DV_{ti} = \beta_{00} + \beta_{01}*NO\_MANIP_i + \beta_{10}*SESS1_{ti} + \beta_{11}*NO\_MANIP_i*SESS1_{ti} + \beta_{20}*PHASE_{ti} + \beta_{21}*NO\_MANIP_i*PHASE_{ti} + \beta_{30}*SXP_{ti} + \beta_{31}*NO\_MANIP_i*SXP_{ti} + r_{0i} + r_{1i}*SESS1_{ti} + r_{2i}*PHASE_{ti} + r_{3i}*SXP_{ti} + e_{ti}$

APPENDIX A
Continued

Taxonomy of Hierarchical Linear Models

Model	Level-1 Model	Level-2 Model	Mixed Model
J	$DV_{ti} = \pi_{0i} + \pi_{1i}*(SESS1_{ti}) + \pi_{2i}*(PHASE_{ti}) + \pi_{3i}*(SXP_{ti}) + e_{ti}$	$\pi_{0i} = \beta_{00} + \beta_{01}*(ISOA_i) + \beta_{02}*(ISOI_i) + r_{0i}$ $\pi_{1i} = \beta_{10} + \beta_{11}*(ISOA_i) + \beta_{12}*(ISOI_i) + r_{1i}$ $\pi_{2i} = \beta_{20} + \beta_{21}*(ISOA_i) + \beta_{22}*(ISOI_i) + r_{2i}$ $\pi_{3i} = \beta_{30} + \beta_{31}*(ISOA_i) + \beta_{32}*(ISOI_i) + r_{3i}$	$DV_{ti} = \beta_{00} + \beta_{01}*ISOA_i + \beta_{02}*ISOI_i + \beta_{10}*SESS1_{ti} + \beta_{11}*ISOA_i*SESS1_{ti} + \beta_{12}*ISOI_i*SESS1_{ti} + \beta_{20}*PHASE_{ti} + \beta_{21}*ISOA_i*PHASE_{ti} + \beta_{22}*ISOI_i*PHASE_{ti} + \beta_{30}*SXP_{ti} + \beta_{31}*ISOA_i*SXP_{ti} + \beta_{32}*ISOI_i*SXP_{ti} + r_{0i} + r_{1i}*SESS1_{ti} + r_{2i}*PHASE_{ti} + r_{3i}*SXP_{ti} + e_{ti}$
K	$DV_{ti} = \pi_{0i} + \pi_{1i}*(SESS1_{ti}) + \pi_{2i}*(PHASE_{ti}) + \pi_{3i}*(SXP_{ti}) + e_{ti}$	$\pi_{0i} = \beta_{00} + \beta_{01}*(GENA_i) + \beta_{02}*(SPEDS_i) + \beta_{03}*(GENI_i) + \beta_{04}*(SPEDI_i) + r_{0i}$ $\pi_{1i} = \beta_{10} + \beta_{11}*(GENA_i) + \beta_{12}*(SPEDA_i) + \beta_{13}*(GENI_i) + \beta_{14}*(SPEDI_i) + r_{1i}$ $\pi_{2i} = \beta_{20} + \beta_{21}*(EX4_i) + \beta_{22}*(EX5_i) + \beta_{23}*(INT4_i) + \beta_{24}*(INT5_i) + r_{2i}$ $\pi_{3i} = \beta_{30} + \beta_{31}*(EX4_i) + \beta_{32}*(EX5_i) + \beta_{33}*(INT4_i) + \beta_{34}*(INT5_i) + r_{3i}$	$DV_{ti} = \beta_{00} + \beta_{01}*GENA_i + \beta_{02}*SPEDA_i + \beta_{03}*GENI_i + \beta_{04}*SPEDI_i + \beta_{10}*SESS1_{ti} + \beta_{11}*GENA_i*SESS1_{ti} + \beta_{12}*SPEDA_i*SESS1_{ti} + \beta_{13}*GENI_i*SESS1_{ti} + \beta_{14}*SPEDI_i*SESS1_{ti} + \beta_{20}*PHASE_{ti} + \beta_{21}*GENA_i*PHASE_{ti} + \beta_{22}*SPEDA_i*PHASE_{ti} + \beta_{23}*GENI_i*PHASE_{ti} + \beta_{24}*SPEDI_i*PHASE_{ti} + \beta_{30}*SXP_{ti} + \beta_{31}*GENA_i*SXP_{ti} + \beta_{32}*SPEDA_i*SXP_{ti} + \beta_{33}*GENI_i*SXP_{ti} + \beta_{34}*SPEDI_i*SXP_{ti} + r_{0i} + r_{1i}*SESS1_{ti} + r_{2i}*PHASE_{ti} + r_{3i}*SXP_{ti} + e_{ti}$
L	$DV_{ti} = \pi_{0i} + \pi_{1i}*(SESS1_{ti}) + \pi_{2i}*(PHASE_{ti}) + \pi_{3i}*(SXP_{ti}) + e_{ti}$	$\pi_{0i} = \beta_{00} + \beta_{01}*(AS1_i) + \beta_{02}*(AS2_i) + \beta_{03}*(IS1_i) + \beta_{04}*(IS2_i) + r_{0i}$ $\pi_{1i} = \beta_{10} + \beta_{11}*(AS1_i) + \beta_{12}*(AS2_i) + \beta_{13}*(IS1_i) + \beta_{14}*(IS2_i) + r_{1i}$ $\pi_{2i} = \beta_{20} + \beta_{21}*(AS1_i) + \beta_{22}*(AS2_i) + \beta_{23}*(IS1_i) + \beta_{24}*(IS2_i) + r_{2i}$ $\pi_{3i} = \beta_{30} + \beta_{31}*(AS1_i) + \beta_{32}*(AS2_i) + \beta_{33}*(IS1_i) + \beta_{34}*(IS2_i) + r_{3i}$	$DV_{ti} = \beta_{00} + \beta_{01}*AS1_i + \beta_{02}*AS2_i + \beta_{03}*IS1_i + \beta_{04}*IS2_i + \beta_{10}*SESS1_{ti} + \beta_{11}*AS1_i*SESS1_{ti} + \beta_{12}*AS2_i*SESS1_{ti} + \beta_{13}*IS1_i*SESS1_{ti} + \beta_{14}*IS2_i*SESS1_{ti} + \beta_{20}*PHASE_{ti} + \beta_{21}*AS1_i*PHASE_{ti} + \beta_{22}*AS2_i*PHASE_{ti} + \beta_{23}*IS1_i*PHASE_{ti} + \beta_{24}*IS2_i*PHASE_{ti} + \beta_{30}*SXP_{ti} + \beta_{31}*AS1_i*SXP_{ti} + \beta_{32}*AS2_i*SXP_{ti} + \beta_{33}*IS1_i*SXP_{ti} + \beta_{34}*IS2_i*SXP_{ti} + r_{0i} + r_{1i}*SESS1_{ti} + DV_{ti} = \pi_{0i} + \pi_{1i}*(SESS1_{ti}) + \pi_{2i}*(PHASE_{ti}) + \pi_{3i}*(SXP_{ti}) + e_{ti}\ r_{2i}*PHASE_{ti} + r_{3i}*SXP_{ti} + e_{ti}$
M	$DV_{ti} = \pi_{0i} + \pi_{1i}*(SESS1_{ti}) + \pi_{2i}*(PHASE_{ti}) + \pi_{3i}*(SXP_{ti}) + e_{ti}$	$\pi_{0i} = \beta_{00} + \beta_{01}*(BEH\_INT_i) + r_{0i}$ $\pi_{1i} = \beta_{10} + \beta_{11}*(BEH\_INT_i) + r_{1i}$ $\pi_{2i} = \beta_{20} + \beta_{21}*(BEH\_INT_i) + r_{2i}$ $\pi_{3i} = \beta_{30} + \beta_{31}*(BEH\_INT_i) + r_{3i}$	$DV_{ti} = \beta_{00} + \beta_{01}*BEH\_INT_i + \beta_{10}*SESS1_{ti} + \beta_{11}*BEH\_INT_i*SESS1_{ti} + \beta_{20}*PHASE_{ti} + \beta_{21}*BEH\_INT_i*PHASE_{ti} + \beta_{30}*SXP_{ti} + \beta_{31}*BEH\_INT_i*SXP_{ti} + r_{0i} + r_{1i}*SESS1_{ti} + r_{2i}*PHASE_{ti} + r_{3i}*SXP_{ti} + e_{ti}$
N	$DV_{ti} = \pi_{0i} + \pi_{1i}*(SESS1_{ti}) + \pi_{2i}*(PHASE_{ti}) + \pi_{3i}*(SXP_{ti}) + e_{ti}$	$\pi_{0i} = \beta_{00} + \beta_{01}*(EVED\_SCO_i) + r_{0i}$ $\pi_{1i} = \beta_{10} + \beta_{11}*(EVED\_SCO_i) + r_{1i}$ $\pi_{2i} = \beta_{20} + \beta_{21}*(EVED\_SCO_i) + r_{2i}$ $\pi_{3i} = \beta_{30} + \beta_{31}*(EVED\_SCO_i) + r_{3i}$	$DV_{ti} = \beta_{00} + \beta_{01}*EVED\_SCO_i + \beta_{10}*SESS1_{ti} + \beta_{11}*EVED\_SCO_i*SESS1_{ti} + \beta_{20}*PHASE_{ti} + \beta_{21}*EVED\_SCO_i*PHASE_{ti} + \beta_{30}*SXP_{ti} + \beta_{31}*EVED\_SCO_i*SXP_{ti} + r_{0i} + r_{1i}*SESS1_{ti} + r_{2i}*PHASE_{ti} + r_{3i}*SXP_{ti} + e_{ti}$
O	$DV_{ti} = \pi_{0i} + \pi_{1i}*(SESS1_{ti}) + \pi_{2i}*(PHASE_{ti}) + \pi_{3i}*(SXP_{ti}) + e_{ti}$	$\pi_{0i} = \beta_{00} + \beta_{01}*(AUTOCORR_i) + r_{0i}$ $\pi_{1i} = \beta_{10} + \beta_{11}*(AUTOCORR_i) + r_{1i}$ $\pi_{2i} = \beta_{20} + \beta_{21}*(AUTOCORR_i) + r_{2i}$ $\pi_{3i} = \beta_{30} + \beta_{31}*(AUTOCORR_i) + r_{3i}$	$DV_{ti} = \beta_{00} + \beta_{01}*AUTOCORR_i + \beta_{10}*SESS1_{ti} + \beta_{11}*AUTOCORR_i*SESS1_{ti} + \beta_{20}*PHASE_{ti} + \beta_{21}*AUTOCORR_i*PHASE_{ti} + \beta_{30}*SXP_{ti} + \beta_{31}*AUTOCORR_i*SXP_{ti} + r_{0i} + r_{1i}*SESS1_{ti} + r_{2i}*PHASE_{ti} + r_{3i}*SXP_{ti} + e_{ti}$

Note. DV is Dependent Variable, SESS1 is Session, SXP is Session by Phase, PXO is Phase by Order, SXPXO is Session by Phase by Order, ADHD is Attention Deficit Hyperactivity Disorder, EBD is Emotional and/or Behavioral Disorders, NONIDENT is Not Identified with a Disability, BEH_TOP is Behavior Topography, COMORBID is Comorbidity, FA us Functional Analysis, NO_MANIP is Descriptive Assessment, ISOA is assessment in an isolated setting, ISOI is intervention in an isolated setting, GENA is Assessment in General Education setting, SPEDA is Assessment in Special Education setting, GENI is intervention in general education setting, SPEDI is intervention in special education setting, AS1 is Researcher Conducted Assessment, AS2 is Teacher Conducted Assessment, INT1 is Researcher Conducted Intervention, INT2 is Teacher Conducted Intervention, BEH_INT is Behavioral Intervention Type, EVED_SCOR is Single-Subject Evidence-Based Practice Score, AUTOCORR is presence or absence of Significant Autocorrelation