Journal of Autism and Developmental Disorders

10.1007/s10803-008-0596-0

Original Paper

Comprehensive Synthesis of Early Intensive Behavioral Interventions for Young Children with Autism Based on the UCLA Young Autism Project Model

Brian Reichow1 and Mark Wolery1

(1) Department of Special Education, Vanderbilt University, Peabody Box 228, Nashville, 37203, TN, USA

Brian Reichow

Email: brian.reichow@vanderbilt.edu

Received: 24 September 2007 Accepted: 15 May 2008 Published online: 6 June 2008

Abstract A 3-part comprehensive synthesis of the early intensive behavioral intervention (EIBI) for young children with autism based on the University of California at Los Angeles Young Autism Project method (Lovaas in Journal of Consulting and Clinical Psychology, 55, 3–9, 1987) is presented. The three components of the synthesis were: (a) descriptive analyses, (b) effect size analyses, and (c) a meta-analysis. The findings suggest EIBI is an effective treatment, on average, for children with autism. The conditions under which this finding applies and the limitations and cautions that must be taken when interpreting the results are discussed within the contextual findings of the moderator analyses conducted in the meta-analysis.

Keywords Autism - Early intensive behavioral intervention - Applied behavior analysis - Lovaas

Introduction

Recently, the Committee on Educational Interventions for Children with Autism of the National Research Council (NRC) reviewed ten comprehensive intervention programs for young children with autism (Lord et al. 2001). Some of these programs were based on applied behavior analysis, which is

a method that has been used to treat children with autism for many years. Recent survey data suggest interventions based on applied behavior analysis are some of the most frequently used interventions in autism (Green et al. 2006; Stahmer et al. 2005).

Many of the programs had supporting empirical evidence, but the NRC did not recommend a single program and cited a need for more research on them (Lord et al. 2001). Instead, consensus guidelines were listed stating children with autism should receive a comprehensive intervention program beginning as soon as they are diagnosed. The program should (a) address the individual's unique deficit areas, (b) use low teacher to student ratios, (c) include a family component, (d) be provided for at least 20–25 h per week, and (e) conduct ongoing assessment and revision of intervention goals and objectives (Lord et al.). Similar guidelines have been recommended by others (Dawson and Osterling 1997; Iovannone et al. 2003; Volkmar et al. 1999) and are generally consistent with recommended practices in early intervention (Sandall et al. 2005).

One comprehensive intervention program reviewed by the NRC (Lord et al. 2001) was early intensive behavioral intervention (EIBI) based on the University of California at Los Angeles Young Autism Project model (UCLA YAP; Lovaas 1981, 1987, 2003). This program was an intensive home-based program using the manual published by Lovaas (1981). The program typically lasted at least 2 years and involved upwards to 40 h of therapy each week. The first results from the program were noteworthy; Lovaas (1987) reported an average difference of 31 points on IQ tests between the treatment and control group, and classified 9 of 19 (47%) participants as having achieved recovery (defined as post-intervention IQ in the normal range—i.e., >85—and successful completion of first grade in a regular education classroom or unassisted placement in a regular education setting). This study, and the claims made by Lovaas (i.e., recovery) caused much debate among researchers. Criticisms focused on methodological limitations including assignment to groups, non-uniform assessment protocol, and selection bias (e.g., Gresham and MacMillan 1998; Mundy 1993; Schopler et al. 1989). Critics often cited the need for additional replications.

Since the 1987 study, replications have occurred including those conducted as part of the National Institute of Mental Health Multi-Site Young Autism Project [MYAP, and independent replications (e.g., Birnbrauer and Leach 1993; Anderson et al. 1987)]. The replications have attempted to address methodological criticisms levied against the original study and have incorporated stronger methods including random assignment to groups (Sallows and Graupner 2005; Smith et al. 2000). Variations of the original intervention protocol also have been examined, including examination of home-based EIBI (Sheinkopf and Siegel 1998), community-based EIBI (Magiati et al. 2007), school-based EIBI (Eikeseth et al. 2007; Eldevik et al. 2006) and parent-managed EIBI (Bibby et al. 2001; Sallows and Graupner 2005).

The purpose of this paper is to provide a comprehensive synthesis of the studies on EIBI. This synthesis includes an examination of the characteristics of the experimental methods, participants, and intervention program (i.e., EIBI), as well an analysis of the effects of EIBI on participants (e.g., outcome data). To accommodate both descriptive and statistical analyses, this synthesis was conducted on multiple levels: (a) descriptive analysis, (b) effect size analyses, and (c) a meta-analysis.

Method

Study Selection

The selection of studies for this review involved seven inclusion criteria: (a) study specified the EIBI was based on the UCLA YAP model by describing the study as a replication of Lovaas (1987), citing intervention techniques and/or curriculum based on one of the Lovaas manuals (Lovaas 1981, 2003), reference to funding from the MYAP, and/or through personal communication with experts who worked with Lovaas on the UCLA YAP or directed MYAP replication sites (J. Wynn, October 9, 2007; M. Amerine-Dickens, March 5, 2007; T. Smith, March 5, 2007, personal communication); (b) participants had diagnoses of autistic disorder, autism spectrum disorder (ASD), pervasive developmental disorder (PDD), or pervasive developmental disorder not otherwise specified (PDD-NOS); (c) participant samples receiving EIBI treatment had a mean chronological age less than 84 months at the beginning of treatment; (d) mean duration of EIBI was greater than or equal to 12 months; (e) at least one child outcome measure was reported; (f) experimental research designs (e.g., pre-test/post-test multiple-group design) or quasi-experimental research designs (i.e., nonequivalent control group design, one-group pre-test/post-test design) were used (Campbell and Stanley 1963); and (g) publication in English in a peer-reviewed journal. A four-step literature search was conducted in the following order: (a) electronic database search, (b) review of references from review articles on comprehensive early intervention programs for children with autism and eligible reports, (c) hand search of selected journals, and (d) expert contact.

Fourteen research reports were located meeting all inclusion criteria and are shown in Table 1. Two reports, Lovaas (1987) and McEachin et al. (1993) used the same participants. It was therefore decided to limit the reports such that each individual (participant of a study) only contributed one result to the synthesis. The Lovaas (1987) report was used because the data were more consistent with other studies. In the Sallows and Graupner (2005) study, two arrangements of EIBI were compared (clinic-coordinated EIBI and parent-coordinated EIBI). In summary, data of 14 samples from 13 research reports were analyzed.1

Table 1 Methodological characteristics of studies

StudyRigorDesignGroup assignmentProcedural fidelityMeasurement constructs by timing of measurement

AdherenceDifferentiationCompetenceIQABLangAPPsyDR

Lovaas (1987) AdequateQuasi-experimental prospective multiple-group comparisonTherapist availabilityIndirect measures, treatment manualNot reportedIndirect measuresPre/post PrePostPrePost

Anderson et al. (1987) WeakQuasi-experimental prospective one-group pre/post designParent selectionDirect measures, treatment manualNot applicable (one-group study)Direct measuresPre/postPre/postPrePostPrePost

Birnbrauer and Leach (1993) WeakQuasi-experimental prospective multiple-group comparisonParent selectionIndirect measures, treatment manualIndirect measuresNot reportedPre/postPre/postPre/post Pre/post

Smith et al. (1997) AdequateQuasi-experimental retrospective multiple-group comparisonTherapist availabilityIndirect measures, treatment manualNot reportedIndirect measuresPre/postPre/postPre Pre/postPost

Sheinkopf and Siegel (1998) WeakQuasi-experimental retrospective multiple-group comparisonParent selectionIndirect measures, treatment manualNot reportedNot reportedPre/post PostPre/post

Smith et al. (2000) StrongExperimental multiple-group comparisonRandom assignmentIndirect measures, treatment manualNot reportedIndirect measuresPre/postPre/postPre/postPostPre/postPost

Bibby et al. (2001) WeakQuasi-experimental retrospective one-group pre/post designParent selectionIndirect measures, treatment manualNot applicable (one-group study)Indirect measuresPre/postPre/postPre Pre/postPost

Boyd and Corley (2001) WeakQuasi-experimental retrospective one-group pre/post designParent selectionIndirect measures, treatment manualNot applicable (one-group study)Direct measures PrePostPre/postPost

Sallows and Graupner (2005) StrongExperimental multiple-group comparisonRandom assignmentIndirect measures, treatment manualIndirect measuresDirect measuresPre/postPre/postPre/postPostPre/postPost

Cohen et al. (2006) StrongQuasi-experimental prospective multiple-group comparisonParent selectionIndirect measures, treatment manualIndirect measuresIndirect measuresPre/postPre/postPre/postPostPrePost

Eldevik et al. (2006) AdequateQuasi-experimental retrospective multiple-group comparisonParent selectionIndirect measures, treatment manualIndirect measuresIndirect measuresPre/postPre/postPre/post Pre/post

Eikeseth et al. (2007) AdequateQuasi-experimental prospective multiple-group comparisonTherapist availabilityIndirect measures, treatment manualIndirect measuresIndirect measuresPre/postPre/postPrePostPre/post

Magiati et al. (2007) AdequateQuasi-experimental prospective multiple-group comparisonParent selectionIndirect measures, treatment manualIndirect measuresNot reportedPre/postPre/postPre/postPostPre/post

AB adaptive behavior, Lang expressive and receptive language, AP academic placement, Psy psychopathology, DR diagnostic recovery

Coding of Study Reports

The study characteristics and outcome data were coded using a manual and forms created for this synthesis. Three study level characteristics (research methods, participant characteristics, and intervention characteristics) were defined and coded to provide information about each study. Outcome data were coded for both samples receiving EIBI and for comparisons between groups receiving EIBI and non-EIBI groups. All coded data (including effect sizes) were obtained directly from the study reports or via contact with a study researcher.

Interobserver agreement (IOA) was assessed on 4 of 14 samples (29%) for the coding of study reports by two independent recorders. IOA was calculated as the product of the quotient of agreements by disagreements and 100. The range of IOA by sample was 85.5–93%. The mean IOA for the four samples reviewed was 91.6%.

Descriptive Analysis

Methodological Characteristics

To assess the influence experimental methods had on study outcomes, five methodological areas were analyzed. First, an overall rating of experimental rigor was obtained using the Evaluative Method for Determining Evidence-Based Practices in Autism (Reichow et al. in press). Second, the study design was categorized (i.e., experimental multiple-group comparison, quasi-experimental prospective multiple-group comparison, quasi-experimental retrospective multiple-group comparison, quasi-experimental prospective one-group pre/post design, or quasi-experimental retrospective one-group pre/post design). Third, the method used for group assignment was categorized into three groups (random assignment, therapist availability, and parent selection).

Fourth, procedural fidelity (Billingsly et al. 1980) was analyzed using the conceptual systems of treatment integrity proposed by Perepletchikova and Kazdin (2005) and Gresham (2005). Perepletchikova and Kazdin (2005) defined three components of treatment integrity. Fidelity of treatment adherence was defined as evidence the characteristics of treatment were delivered consistently as planned across and within participants of a sample. Treatment differentiation was defined as evidence the groups of a comparative study received different levels of the treatment package. Therapist competence was defined as evidence of therapist training and/or evaluation of therapist performance. In a response to these components, Gresham (2005) outlined three methods of measuring treatment integrity: (a) direct measures, (b) indirect measures, and (c) manualized treatments. The final methodological characteristic was the measures used, which was categorized into six constructs (IQ, adaptive behavior, language, academic placement, psychopathology, and diagnostic reclassification).

Participant Characteristics

Participant characteristics were assessed by examining the pre-treatment assessments on six variables: (a) diagnosis, (b) chronological age, (c) IQ, (d) adaptive behavior, (e) language, and (f) other treatments received. These data were used to illustrate differences between samples and as moderator variables for the meta-analysis.

Intervention Characteristics

Nine intervention characteristics were identified for this review. Three intervention characteristics pertained to the intensity of the intervention. Intervention density was defined as the total number of hours per week participants received EIBI. Intervention duration was defined as the total number of months each participant received EIBI. The total hours of therapy was calculated by multiplying the product of intervention density and duration by 4.3 (converter for months to weeks). Not all studies reported the mean intervention density and/or duration. When the mean data for density and/or duration were not provided, an estimated value was determined from information in the study report and used for all subsequent analyses.

Three intervention characteristics described the organization of intervention services. The model of supervisor training was a dichotomous variable; studies were either categorized as being consistent with the UCLA/MYAP training protocol, including an internship at an affiliated clinic site (i.e., UCLA or MYAP), or studies were categorized as using other training models (e.g., inservice, on-the-job, workshop-based). The second organizational intervention characteristic categorized the type of service coordination model as being clinic-coordinated, community-coordinated, or parent-coordinated. Parental role was defined by the type of involvement expected for each participant's parents (usually mother). These included conducting therapy, service-coordination, and assisting therapists.

The remaining three intervention characteristics describe aspects of the EIBI therapy. The educational and/or training qualifications of therapist were categorized as parent, undergraduate college student, lay person, or paraprofessional. The location of therapy was coded as the location intervention occurred across the entire intervention period and included three categories (home, school, community). Finally, the use of physical aversives was recorded as occurring, not occurring, or not reported for each sample.

Outcome Data

Descriptive analyses were conducted on constructs with no pre-intervention assessment (academic placement, diagnostic reclassification) and for constructs using many different measures (psychopathology). Because the calculation of an effect size was not appropriate for these constructs, they were analyzed using descriptive statistics. These analyses were conducted on the sample data,2 thus the results reflect the changes within a sample without reference to a control group. The data for placement were analyzed by reporting the range of the percentage of participants from each sample in regular education classrooms and other educational settings (e.g., special education settings, aphasic classrooms). The data for psychopathology were analyzed by comparing the mean scores of the pre- and post-intervention assessments for each sample, which were then categorized by the type of change. The data for diagnostic classification were analyzed by reporting the range of the percentage of participants meeting Lovaas' (1987) criteria of recovery (i.e., post-intervention IQ in the normal range—i.e., greater than 85—and successful completion of first grade in a regular education classroom or unassisted placement in a regular education setting) for each sample.

Effect Size Analyses

Effect sizes were calculated for the outcome data from the constructs of IQ, adaptive behavior, expressive language, and receptive language. Two types of effect sizes were used: The standardized

mean change effect size and the standardized mean difference effect size. The formulae for these are shown in Table 2. Three steps were taken to help ensure the most conservative effect sizes were calculated. First, effect sizes were calculated only when the data necessary for its calculation were available. If a sample or study was missing the necessary data for the calculation of an effect size, no effect size was calculated for that study. Hence, no data were extrapolated or interpolated for the calculation of effect sizes. Second, Hedge's g (Hedges and Olkin 1985) was used as the effect size metric, which calculates a more conservative (i.e., smaller) estimate of the effect size than Glass' Δ or Cohen's d (Grissom and Kim 2005). Finally, because effect sizes based on small samples are known to be biased (Lipsey and Wilson 2001), effect sizes were multiplied by the small sample correction factor (Hedges and Olkin 1985).

Table 2 Formulas used in the statistical analyses

FormulaEquationWhere

Standardized mean change effect size with small sample adjustmentg c = d {1 - [3/(4 × df - 1)]} d = (Y 2 − Y 1)/s p 2

df = Degrees of freedom

Y 1 = pre-treatment mean

Y 2 = post-treatment mean

s p 2 = √[(n 1 − 1)s 1 2 + (n 2 − 1)s2 2)/n 1 + n 2 − 2]

n 1 = Number of participants at pre-treatment

n 2 = Number of participants at post-treatment

s 1 2 = Pre-treatment variance

s 2 2 = Post-treatment variance

Standardized mean difference effect size with small sample adjustmentg d = d {1 − [3/(4 × df − 1)]} d = (Y 2 − Y 1)/s p 2

df = Degrees of freedom

Y 1 = mean for comparison group

Y 2 = mean for EIBI group

s p 2 = √[(n 1 − 1)s 1 2 + (n 2 − 1)s 2 2)/n 1 + n 2 − 2]

n 1 = Number of participants in comparison group

n 2 = Number of participants in EIBI group

s 1 2 = Variance of comparison group

s 22 = Variance of EIBI group

Small sample correction1 − [3/(4 × df − 1)] df = Degrees of freedom

Mean effect size = ΣESi × w i/Σw i ESi = g c for studies i = 1 to k

w i = Inverse variance weight = 1/SE2

Q-statistic Q = Σw i (ESi −)2 w i = Inverse variance weight = 1/SE2

ESi = g c for studies i = 1to k

 = standardized mean effect size

Proportion of variance accounted for by between-study varianceI 2 = Q/(df − 1)/(Q/df) Q = Σw i (ESi −)2

df = Degrees of freedom = (n − 1)

The first effect size analyses were calculated using the standardized mean change effect size and examined the difference between the average gains made by distinct samples. This comparison showed the absolute difference within a sample without regard to a comparison or control group. For these analyses, the effect sizes were analyzed with reference to the research report rigor rating (i.e., strong, adequate, and weak) of the study containing the sample.

For the ten studies using between-group designs, the standardized mean difference effect size (g d) was used (shown in Table 2). This effect size showed the magnitude of difference between the group receiving EIBI and the comparison group. For the analyses of the comparative studies, effect sizes for each construct were analyzed with reference to the characteristics of the comparison group. Three types of comparison groups were used across studies. Two studies (Lovaas 1987; Smith et al. 1997) compared intensity of behavioral intervention (i.e., high intensity vs. low intensity). Six studies (Birnbrauer and Leach