

Systematic Review of Early Intensive Behavioral Interventions for Children With Autism

Patricia Howlin and Iliana Magiati

Institute of Psychiatry, King's College (London, UK)

Tony Charman

University College, London, Institute of Child Health

Abstract

Recent reviews highlight limitations in the evidence base for early interventions for children with autism. We conducted a systematic review of controlled studies of early intensive behavioral interventions (EIBI) for young children with autism. Eleven studies met inclusion criteria (including two randomized controlled trials). At group level, EIBI resulted in improved outcomes (primarily measured by IQ) compared to comparison groups. At an individual level, however, there was considerable variability in outcome, with some evidence that initial IQ (but not age) was related to progress. This review provides evidence for the effectiveness of EIBI for some, but not all, preschool children with autism.

DOI: 10.1352/2009.114:23–41

Over the past 2 decades, there has been increasing interest in developing effective interventions for young children with autism spectrum disorder. In addition to the many largely untested “alternative” therapies (from acupuncture to Zinc supplements; see Research Autism, 2007), a wide range of more soundly based psychoeducational programs has been employed. These tend to incorporate a mix of behavioral, developmental, and educational approaches (for reviews, see National Research Council, 2001; Scottish Intercollegiate Guidelines Network [SIGN], 2007), and although methods vary, the general goal of such programs is to enhance cognitive, communication, and social skills while minimizing autistic symptomatology and other problem behaviors. Three main strands of intervention have been the focus in the majority of studies conducted to date: programs that have a specific focus on communication; those in which developmental/educational strategies have been employed, and those with a particular emphasis on the use of behavioral principles to improve learning and behavior.

In terms of communication-focused interven-

tions, programs such as the Picture Exchange Communication System (Bondy & Frost, 1998) and other alternative/augmentative communication systems were designed in an attempt to provide a communication modality for children who have no spoken language (see Howlin, 2006). Other programs, designed for both nonverbal and verbal children, use approaches informed by psycholinguistic theory to target the early interactions between parents and newly diagnosed children in order to enhance nonverbal and verbal communication. These include parent communication training approaches (Aldred, Green, & Adams, 2004; Drew et al, 2002); More Than Words (Sussman, 1999), Early Bird (Shields, 2001), and Pre-Linguistic Milieu Therapy (Yoder & Stone, 2006). In addition, there are numerous developmental and educational programs that combine aspects of developmental, educational, and behavioral approaches, including Daily Life Therapy (Quill, Gurry, & Larkin, 1989); the Denver Model (Rogers et al., 2006); the Douglass Developmental Disabilities Center program (Handleman, Harris, Arnold, Gordon, & Cohen, 2006); Floor Time/De-

velopmental, Individual Difference Relationship model (Greenspan & Wieder, 2003); Son-Rise (Kaufman, 1994); TEACCH (Treatment and Education of Autistic and Communication Handicapped Children; Schopler, 1997), and many more.

Although all these programs employ somewhat different methodologies, they also have important elements in common; for example, techniques developed from learning theory (Skinner, 1953) are essential components of most. Operant approaches, as exemplified in the Applied Behavior Analysis (ABA) approach (Dunlap, Kern-Dunlap, Clark, & Robbins, 1991) are particularly fundamental to behavioral techniques such as Pivotal Response Training (Schreibman & Koegel, 2005), Discrete Trial Training (Maurice, Green, & Luce, 1996) and Verbal Behavior (Barbera & Rasmussen, 2007) that form part of many modern-day early intervention programs for children with autism.

Recently, reviewers of intervention programs for children with autism have been critical of the standards of research in this field, pointing to various methodological limitations and failures of experimental design (Levy, Kim, & Olive, 2006; Scottish Intercollegiate Guidelines, 2007; Wheeler, Blaggett, Fox, & Blevins, 2006). However, there is also widespread agreement that interventions based on ABA, particularly those involving home therapy and beginning in the preschool years, have been most comprehensively studied and, as a consequence, have the best established evidence base (Lord et al., 2005; National Research Council, 2001; Smith, Donahoe, & Davis, 2006). Among the most thoroughly evaluated are programs involving early intensive home-based behavioral intervention (EIBI). The EIBI approach underlies the UCLA Young Autism Project, originally developed by Lovaas and his colleagues (1981). Other more recently developed EIBI programs include those described by Maurice, Green, and Luce (1996), Maurice, Green, and Foxx (2001), and Handleman et al. (2006).

Lovaas was among the earliest researchers to demonstrate the effectiveness of behavioral interventions for children with autism and was also one of the first to stress the importance of parental involvement in therapy. Proponents of EIBI recommend that therapy should begin as early as possible, preferably before the age of 3 years, should take place for approximately 40 hours per week and last at least 2 years. The programs are highly prescriptive, with detailed manuals provid-

ed to guide and monitor treatment. Learning sessions are provided in a one-to-one discrete trial format, focusing on the systematic teaching of measurable behavioral units, repetitive practice, and structured presentation of tasks from the most simple to the more complex. It has now been 2 decades since Lovaas (1987) published his original paper on the effectiveness of EIBI, reporting that preschool children involved in 1:1 therapy for 40 or more hours a week, over at least 2 years, showed major gains (of up to 30 points) in IQ; rates of integration in mainstream school increased significantly; and 47% of children were described as attaining “normal intellectual and educational functioning (p. 3).” A subsequent follow-up (McEachin, Smith, & Lovaas, 1993) also confirmed that the gains were maintained into early adolescence.

Although in some quarters this study was hailed as “a triumph of behavioral science and behaviorally scientific clinical application” (Baer, 1993, p. 373), critics highlighted problems in experimental design (particularly with regard to non-random assignment and intelligence test–retest measures) and the practical and financial difficulties of replication given the intensity of the program. It was claims of “normal functioning,” however, that gave rise to most contention (Mesibov, 1993; Mundy, 1993). This was partly due to the limited range of assessments employed on which to base judgments of “normality” but also because the figure of 47% showing “normal functioning” (McEachin et al., 1993, p. 367) represented only 9 out of 19 children. Smith et al. (2001) noted that the 9 children classified as having normal functioning accounted for most of the gains in IQ. This subgroup showed an average IQ increase of 37 points compared with an average increase of 3 points for the other 10 children. Moreover, Shea (2004) pointed out that the published data from the studies of Lovaas (1987) and McEachin et al. (1993) indicated that on the instruments used, only one child had test scores that were all within the average range.

The extent of involvement required by parents and therapists to follow the UCLA EIBI program for the recommended number of hours has restricted opportunities for replication, and in the few studies that have been conducted, researchers have generally used a modified version of this model. Nevertheless, a number of replications (or at least partial replications) of the original study has now been undertaken, and although many do

not report gains of the magnitude reported by Lovaas (1987) or McEachin et al. (1993), the results have generally been positive.

Although there have been previous reviews of early interventions for children with autism (Lord et al., 2005; Rogers, 1998; Smith, Eikeseth, Klevstrand, & Lovaas, 2007), there have been many new controlled studies in recent years (Rogers' 1998 review, for example, appeared before 8 of the 11 studies reviewed here were published). Thus, an updated review is required in order to take account of the new evidence reported. In this systematic review, we examined the findings from controlled EIBI studies published in peer-reviewed journals. Our goal was to examine what conclusions can be drawn from research conducted to date and to identify the challenges for future researchers.

Method

The following databases were searched for peer-reviewed papers from 1985 to May 2007: MEDLINE, EMBASE, Cochrane, PSYCHINFO, CINAHL, ERIC, using the terms *autism/autistic/autism spectrum disorder* and *intervention/early intervention/behavioral intervention/EIBI/ABA*. Additional sources of information were the United Kingdom National Autistic Society Research Autism website; international reviews published by the New York State Department of Health Early Intervention Program (1999); the state of Maine (Maine Administrators of Services for Children With Disabilities (Maine, 1999); National Research Council (2001); National Initiative for Autism: Screening and Assessment (2003); University of Sydney (Roberts, 2004); New Zealand Ministry of Health (2006); and National Health Service (NHS) Scotland (Scottish Intercollegiate Guidelines Network, 2007). The inclusion criteria for the review were (a) case-control comparison study (not necessarily randomized design) of EIBI for children with autism; (b) UCLA-affiliated (home- or center-based) or other home-based EIBI program largely based on the UCLA model (not necessarily for 40 hours per week); (c) minimum of 10 participants in EIBI group; (d) age at start of treatment less than 6 years; (e) intervention minimum 12 hours per week; (e) duration minimum 12 months; (f) adequate data on IQ or other standard measures to allow calculation of extent of gains.

Results

In total, we identified 641 studies that met the search term combinations. We excluded studies in which the title indicated clearly that the study was of a single case or a case series without a comparison group and studies with a focus on a specific behavior(s) (e.g., sleeping problems, challenging behavior), the use of pharmacological interventions or other non-ABA programs, or on outcomes for/views of therapists or parents. Abstracts of review papers were checked, and if these were reviews of ABA/EIBI with participants who had autism their reference lists were further scrutinized for possible papers to include. We initially identified 37 papers that appeared to meet criteria for inclusion; 8 of these were overviews of research in the area or commentaries on previously published studies; 5 were non-EIBI or a mix of EIBI and other behavioral interventions; 11 had no reported IQs and/or no comparison group and/or had fewer than 10 participants. Thirteen studies (indicated in boldface in the *References*) met all inclusion criteria; 2 of them (Eikeseth, Smith, Jahr, & Eldevik, 2007; McEachin et al., 1993) were extensions of previous reports (Eikeseth, Smith, Jahr, & Eldevik, 2002; Lovaas, 1987, respectively). These are presented as single outcome reports (see Table 1). Only one study (Smith, Groen, & Wynn, 2000) was a fully randomized control trial. Because data were not available for individual participants, the following summaries/analyses are based on comparisons of the published group means. All but 4 studies (Eikeseth et al., 2007, 2006; Magiati, Charman, & Howlin, 2007; Remington et al., 2007) were carried out in the United States.

Characteristics of Participants and Therapeutic Interventions

Diagnostic measures varied across studies, with most relying on clinical judgments based on *Diagnostic and Statistical Manual of Mental Disorders/International Classification of Diseases* criteria. Six studies (Cohen, Amerine-Dickens, & Smith, 2006; Eikeseth et al., 2002, 2007; Eldevik, Eikeseth, Jahr, & Smith, 2006; Magiati et al., 2007; Remington et al., 2007; Sallows, & Graupner, 2005) used the Autism Diagnostic Interview-Revised—ADI-R (Lord, Rutter, & le Couteur, 1994) to confirm diagnosis. Reported diagnoses included children with autism, autism spectrum disorder, and pervasive developmental disorder.

Table 1. Summary of Intervention and Participant Characteristics

Study	Diagnostic label ^a	n ^b	Mean ages ^c at Time 1	Type of study	EIBI condition	Comparison (C) condition
Lovaas, 1987; McEachin et al., 1993 (USA)	DSM/Clinical diagnosis of autism	EIBI = 19 C1 = 19 C2 = 21	EIBI = 35 C1 = 41 C2 < 42	Case control trial	UCLA- EIBI	C1: less intensive therapy, C2: minimum therapy
Smith et al., 1997 (USA)	DSM: MR and autism	EIBI = 11 C = 10	EIBI = 36 C = 38	Retrospective case-control study	UCLA-EIBI multisite	Minimum therapy
Sheinkopf & Siegel, 1998 (USA)	DSM: autism; PDD	EIBI = 11 C = 11	EIBI = 34 C = 35	Retrospective case-control trial	Community-based UCLA model	Standard schooling
Smith, Groen, & Wynn, 2000 (USA)	Clinic diagnosis: autism; PDD	EIBI = 15 C = 13	EIBI = 36 C = 36	Randomized control trial	UCLA-EIBI multisite	Parent-training/public school
Eikeseth et al., 2002/2007 (Norway)	ADI-R autism	EIBI = 13 C = 12	EIBI = 66 C = 65	Case-control trial	School-based UCLA model	Eclectic education
Sallows & Graupner, 2005 (USA)	ADI-R autism	EIBI = 13 C = 10	EIBI = 35 C = 37	Randomized allocation to clinic or parent-run group	UCLA-EIBI multisite	Parent-directed EIBI
Howard et al., 2005 (USA)	DSM (with severity details) autism; PDD	EIBI = 29 C1 = 16 C2 = 16	EIBI = 31 C1 = 39 C2 = 35	Case-control trial	Community-based EIBI	C1: 6 autism specific ed. C2: 16 generic special ed.
Cohen et al., 2006 (USA)	ADI-R autism; PDD	EIBI = 21 C = 21	EIBI = 30 C = 33 ^d	Case-control trial	UCLA -EIBI multisite	Eclectic education
Eidevik et al., 2006 (Norway)	Most ADI-R autism + LD	EIBI = 13 C = 15	EIBI = 53 C = 49	Retrospective case-control trial	School/home-based UCLA model	Eclectic education
Magiati et al., 2007 (UK)	ADI-R autism; ASD	EIBI = 28 C = 16	EIBI = 38 C = 42	Case-control trial	Community-based EIBI	Autism specific nursery
Remington et al., 2007 (UK)	ADI-R autism	EIBI = 23 C = 21	EIBI = 36 C = 38	Case-control trial	Community-based EIBI	Mix of therapies, none 1:1

^aDSM = *Diagnostic and Statistical Manual of Mental Disorders*; PDD = pervasive developmental disorder; ADI-R = Autism Diagnostic Interview-Revised; MR = mental retardation. ^bEIBI = early intensive behavioral interventions; C = comparison group (C1 = Group 1; C2 = Group 2 in studies with more than 1 comparison group). ^cIn months. ^dOnly age at diagnosis given.

Smith et al. (1997) and Eldevik et al. (2006) focused specifically on children with intellectual impairments; the remainder included children with mean IQs ranging from 50 to 80+.

The mean number of EIBI children per study was 17.8 ($SD = 6.6$, range = 11 to 29); the mean number of comparison cases was 14.7 ($SD = 4.1$, range = 10 to 21). There was no significant difference in the ages of EIBI and comparison children at baseline; the mean age of the EIBI children was 40 months, with a range of 31 to 66. Comparison children's mean age was 42 months, with a range of 35 to 65. However, it was occasionally difficult to establish accurately from the published data the age at which intervention actually began, with some researchers (e.g., Cohen et al., 2006) reporting only the ages at which children were diagnosed or first assessed rather than the precise age at which intervention started. In the majority of interventions, investigators claimed to follow the behavioral program developed by Lovaas and colleagues as part of the UCLA Young Autism Project; 5 were part of the UCLA multisite replication group; 2 were school-based but affiliated with the UCLA program; 4 were community-based programs. Of these, Sheinkopf and Siegel (1998), according to parental report, followed the UCLA model; Howard, Sparkman, Cohen, Green, and Stanislaw (2005) and Remington et al. (2007) reported that therapists had been trained in a number of ABA/EIBI techniques, including the manualized programs of Maurice et al. (1996, 2001), discrete trial and incidental teaching (Anderson & Romanczyk, 1999), Natural Environment Training (Sundberg & Partington, 1999), and Verbal Behavior (Partington & Sundberg, 1998); Magiati et al. (2007) included at least some consultants trained at UCLA.

The comparison conditions varied from intensive, parent-directed (as opposed to clinic-directed) intervention (Sallows & Graupner, 2005) to reduced intensity of EIBI program (Lovaas, 1987; Smith et al., 1997); eclectic, public schooling (Cohen et al., 2006; Eikeseth et al., 2002, 2006; Eldevik et al., 2006; Sheinkopf & Siegel, 1998); specialist autism school (Howard et al., 2005; Magiati et al., 2007) or a mixture of different interventions (Remington et al., 2007). Two studies (Howard et al., 2005; Lovaas, 1987) involved two separate comparison groups.

Duration and Intensity of Interventions

Table 2 provides information on the approximate duration of intervention and the ages at

which participants were followed-up. Not all papers provided clear information on how long children were involved in the program, and in some cases children had been enrolled in the therapy some time prior to the initial "baseline" assessments. For some studies we were not able to establish the length of time children were actually involved in intervention. Lovaas (1987) stated that the EIBI children received between 2 and 6 years or more of intervention. Smith et al. (1997) reported that the EIBI children received a minimum of 2 years of therapy; control participants received less than 2 years. As can be seen in Table 2, of the remaining 9 studies, 3 involved interventions of less than 24 months; 5 interventions were 24 to 36 months, and in 1 (Sallows & Graupner, 2005), interventions lasted 4 years. The mean intervention time in these 9 studies for both EIBI and comparison groups was (as far as could be estimated) 27.4 months ($SD = 10.7$, range 14 to 48 months).

Actual hours of intervention proved even more difficult to estimate. Only a minority of researchers (Magiati et al., 2007; Remington et al., 2007; Smith, Groen, & Wynn, 2000) provided their own figures, based on parental or therapists' reports, of the total number of hours in therapy. Most gave only a general indication of hours of intervention (e.g., minimum number of hours per week); several indicated that time in therapy declined after the first year, without providing information on the degree of reduction. In others (e.g., Lovaas, 1987), it was clear that individual variation was considerable, with some children in this, the original Lovaas study, receiving over 6 years and more than 14,000 hours of therapy, while others in the same study received 40 hours of therapy per week for 2 years (i.e., approximate estimation of 4,000 hours). Although it was not possible, with one exception (Smith et al., 2000: $M = 2,158$ hours, $SD = 1,305$, range = 1,142 to 5,452), to derive an accurate estimate of total number of hours in therapy over the duration of treatment, most investigators gave an indication of approximate hours per week, at least during the most intensive phase of intervention. On the basis of the information provided, we estimated that EIBI children received significantly more hours of intervention per week than did controls (EIBI $M = 29.8$ hours per week; $SD = 7.6$, range 12.5 to 40; controls $M = 19.1$ hours, $SD = 8.6$, range ≤ 10.0 to 31.5), $t = 3.03$, $p = .007$.

Despite the limited quality of the available

Early intensive behavioral interventions

P. Howlin, I. Magiati, and T. Charman

Table 2. Summary of Treatment Duration and Intensity

Study	Treatment duration	Follow-up (FU) conducted/Mean age	Approximate number of hours per week	
			EIBI group	Comparison group
Lovaas, 1987; McEachin et al., 1993	From 2 to 6+ years	FU after 2+ and 8+ years (at approx. age 7 years & 11.5 years)	40 hours for 2 years, then decreasing	Less than 10 hours per week both comparison groups
Smith et al., 1997	EIBI ^a ≥ 2 years; C ^b ≤ 2 years	FU after 2.5 to 3 years Age: EIBI 71 months; C 64 months (but some children aged 9–10 years)	30 hours per week	≤10 hours per week
Sheinkopf & Siegel, 1998	EIBI 16 months; C 18 months	FU: EIBI after 20 months; C after 18 months Age: EIBI & C 54 m	27 hours per week	11 hours per week
Smith, Groen, & Wynn, 2000	2–3 years	FU after approx. 5 years Age: EIBI 94 months; C 92 months	25 hours per week 1st year; then reducing	15–20 hours per week (parent and school based)
Eikeseth et al., 2002, 2007	EIBI 31 months; C 33 months	FU after 12 months & 2.5 to 3 years Age: EIBI & C, 98 months	28 hours per week 1st year; 18 hours 2nd year	29 hours per week 1st year; 16 hours 2nd year
Sallows & Graupner, 2005	48 months	FU after 48 months Age: EIBI & C 83 months	39 hours per week 1st year; 37 hours per week, 2nd year	32 hours per week 1st year; 31 hours per week, 2nd year
Howard et al., 2005	13–14 months	FU after 14 months Age: EIBI 46 months; C1 51 months; C2 49 months	25–40 hours per week	C1 25–30 hours per week; C2 mean 15 hours per week
Cohen et al., 2006	36 months (but some EIBI groups had up to 400 hours EIBI before intake)	FU after 36 months Age at FU unclear	35–40 hours per week	15–25 hours per week
Eldevik et al., 2006	EIBI 20 months; C 21 months	FU after 20–21 months Age: EIBI 73 months; C 70 months	12.5 hours per week	12 hours per week
Magiati et al., 2007	24 months	FU after 26 months Age: EIBI 64 months; C 68 months	32 hours per week	26 hours per week
Remington et al., 2007	24 months	FU after 12 & 24 months Age: EIBI & C 4 & 5 years	26 hours per week in EIBI + 22/23 children in school for 9–13 hours per week	15–17 hours per week in school 1st year; 14–22 hours per week 2nd year

^aEarly intensive behavioral interventions. ^bComparison group (C1 = Group 1; C2 = Group 2 in studies with more than one comparison group).

information, we also believed it important to attempt an approximate estimate of total hours in therapy in order to assess the potential impact of intensity on outcome. This estimate was derived by multiplying Approximate Weeks of Intervention \times Approximate Hours per Week (allowing for any reduction in hours reported by authors after 1 or 2 years in therapy). The resulting figures were as follows: EIBI group mean was 3,353 hours ($SD = 1,960$, range = 1,000 to 7,100); comparison group mean was 1,980 hours ($SD = 1,529$, range 836 to 5,952). Although the intensity of intervention was considerably less in the comparison groups than for the EIBI children, this difference fell short of statistical significance, $t = 1.78$, $p = .09$, reflecting the very wide range of hours in both groups.

Baseline and Outcome Data

Just as the studies varied with respect to the characteristics of children involved and the duration and intensity of intervention, there was considerable variation in the methods used to assess outcome. First, not all researchers presented outcome data at the time intervention actually ceased; they instead provided this information as measured at follow-up sometime later. Thus, whereas the mean duration of the intervention program itself was 27 months for both EIBI children and controls (at least in the 9 studies for which a calculation of duration was possible), the mean time between baseline assessment and follow-up was 39.2 months ($SD = 30.1$, range = 12 to >121) for EIBI children and 34.5 months ($SD = 19.2$, range = 12 to 79) for controls. Although the group difference in time from cessation of treatment to follow-up was not significant, $t = .45$, $p = .66$, even a relatively small time lag following the end of intervention further complicated any assessment of the association between duration of therapy and degree of improvement.

Differences in the outcome measures used are a further factor affecting comparisons among studies. Although data on IQ changes over time were provided in all 11 studies, information on the instruments used was generally sparse. The tests used varied, not only from study to study, but from child to child and from baseline to follow-up within the same study, and this variability in assessments over time can result in spurious conclusions about the extent of cognitive improvement (Magiati & Howlin, 2001). In 6 studies data for verbal and nonverbal IQ were reported

separately and in 2, investigators reported changes in mental age (MA). Remington et al. (2007) also calculated the Reliable Change Index (Jacobson & Truax, 1991) in order to establish, with 95% certainty, that the level of IQ change was not due to measurement unreliability/score variability.

The Vineland Adaptive Behavior Scales—VABS (Sparrow, Balla, & Cicchetti, 1984) have been used in the more recent studies (Cohen et al., 2006; Eikeseth et al., 2002, 2007; Eldevik et al., 2006; Howard et al., 2005; Magiati et al., 2007; Remington et al., 2007; Sallows & Graupner, 2005; Smith et al., 1997). Some authors presented standard scores; others, age equivalents or raw scores (e.g., Remington et al., 2007). The latter, in particular, are difficult to interpret with respect to the clinical meaning of any changes found.

Language data have become better standardized over time, with most investigators providing information on expressive and language skills as assessed by the Reynell scales, although not all used the same versions of the Reynell (Cohen et al., 2006; Eikeseth et al., 2002, 2007; Eldevik et al., 2006; Howard et al., 2005; Magiati et al., 2007; Remington et al., 2007; Sallows & Graupner, 2005; Smith et al., 1997).

Again, although some researchers presented standard scores, others cited age equivalents. Magiati et al. (2007) reported raw scores because so few children were able to score above basal and, for the same reason, Remington et al. (2007) simply recorded the number of children able to achieve any score. Few investigators provided adequate information on how issues relating to test basals for standard scores and age equivalents were dealt with. Raw scores are, of course, acceptable when children are so far below basal that age equivalents and standard scores cannot be calculated (Charman, 2004), but when only raw scores are provided by some authors, and only standard scores or age equivalents by others, comparisons among studies become extremely difficult. Including data on age equivalents, standard scores, and raw scores would be far more informative and aid the interpretation of any significant group differences. Now that many journals provide access to online-only appendices, this comprehensive reporting of data should be possible in the future.

Other outcome data vary to such an extent that comparisons between studies are not feasible. Some (e.g., Lovaas, 1987, Sallows & Graupner, 2005) provided information on personality tests

(e.g., the Personality Inventory for Children: Wirt, Lachar, Klinedinst, & Seat, 1984; Smith et al., 1997); several authors assessed levels of problem behaviors using a range of different measures, including maladaptive behavior scales, behavioral observations, or parental reports (Lovaas, 1987; Magiati et al., 2007; Remington et al., 2007; Smith et al., 1997; Smith et al., 2000). In some later studies (Magiati et al., 2007; Sallows & Graupner, 2005), researchers also used the ADI-R to monitor changes in the severity of autism symptomatology over time (although the Autism Diagnostic Observation Schedule–Generic–ADOS-G (Lord et al., 2000) might be a better potential outcome measure of autism severity (see Aldred et al., 2004; Howlin, Gordon, Pasco, Wade, & Charman, 2007). Most studies provide some limited information on the school status of participants following intervention (e.g., whether unsupported in mainstream school).

Table 3 presents the information on changes in IQ over time as provided in all studies. The IQ data indicate that in the majority of studies, there were no significant differences between EIBI and comparison children at baseline; in one study (Magiati et al., 2007), where there was a group difference in IQ (though not MA), this was controlled for in the analysis of change. Comparison of group scores on other tests for which the majority of studies provide information (Vineland Composite; Language Comprehension, and Expression) also indicates no significant group differences at intake. In the studies reviewed different strategies for analysis were used. The most common analysis was a simple *t* test, ANOVA, or Mann Whitney *U* test on outcome scores (Eikeseth et al., 2002; Lovaas, 1987; Sheinkopf & Siegel, 1998; Smith et al., 1997; Smith et al., 2000), including for Smith et al. and Eikeseth et al., one-tailed *t* tests). Eikeseth et al. (2002, 2007) and Eldvik et al. analyzed change scores.

In several of the more recent studies, researchers reported either simple ANCOVAs on outcome scores (Howard et al., 2006; Remington et al., 2007), repeated measures ANOVA (Sallows & Graupner, 2005), or ANCOVA (Cohen et al., 2006; Remington et al., 2007), partly reflecting their larger sample size that enabled them to meet more of the statistical assumptions that such tests require. Although these different analyses do not allow a uniform estimate of effect size to be calculated from the available group scores, in order to provide a simple estimate of effect size for IQ,

we calculated EIBI versus comparison group mean differences divided by the *SD* of the two groups combined (Cohen's *d*) (see Table 3).

Data on school placement at follow-up were provided in the majority of studies (see Table 4). Overall differences, based on study group means for IQ, VABS, and language test standard scores are reported in Table 5 (some data were extrapolated from the published MA scores).

At final follow-up, there were significant group differences in IQ in 9 of the 11 reports. Of the remaining 2 studies, Magiati et al. (2007) reported no differences between the EIBI and comparison groups. Sallows and Graupner (2005) did not analyze for Group \times Time differences; instead, they reported IQ improvements in both groups while noting considerable variability in both groups at outcome. Estimated effect sizes were moderate, >0.60 , to large, >0.80 , in the majority of studies. However, in other studies showing significant results (e.g., Eldevik et al., 2005), the effect size was small, 0.27. The significant group difference in this case is due in part to the analysis of change scores (Time 1 to Time 2) and the fact that although the group mean score increased somewhat for the EIBI group, it decreased in the comparison group. As is evident from the summary data shown in Table 5, there was a substantial mean increase in IQs in the EIBI groups over time. There were also positive changes, of a similar magnitude, in language scores. The improvement in VABS scores, although significant in the summary EIBI versus comparison group analysis, was relatively small (average 5 standard score points).

Several additional points should be noted. First, the extent of change reported in the different studies is considerable, as is clear from Table 5. For example, for IQ, investigators in 3 studies (Eldevik et al., 2006; Magiati et al., 2007; Smith et al., 1997) reported changes of less than 10 points; 2 (Remington et al., 2007; Sheinkopf & Siegel, 1998), a 10 to 20 points increase; 4 (Cohen et al., 2006; Eikeseth et al., 2002, 2007; Sallows & Graupner, 2005; Sheinkopf & Siegel, 1998), a 21 to 30 points increase; and in 2 studies (Howard et al., 2005; Lovaas, 1987), an increase of 31 points. The data on language and VABS scores show similar variation. Also of note is the variation of the change in mean scores of the comparison groups. Although in some studies the mean comparison group IQ increased between baseline and outcome (albeit not as much on average as

Table 3. Summary of IQs at Baseline (Time 1) and Follow-Up (Time 2) by Groups

Study	EIBI IQ baseline		EIBI IQ follow-up (FU)		C IQ FU T2 Mean (SD)	T2 (T3) EIBI vs. C effect size estimate	Statistical test
	1 Mean (SD)	C IQ baseline T1 Mean (SD)	EIBI IQ at subse- quent FU = 84.5 (SD 32.4)	T2 Mean (SD)			
Lovaas, 1987; McEachin et al., 1993	53.0 [30-82] ^a	46.0 [30-80]	83.3 ^b (IQ at subse- quent FU = 84.5 (SD 32.4))	C1 52.2 ^b (IQ at subse- quent FU = 54.9, SD 29.1) C2 57.5	T2 not calculated ^d T3 = 0.96	ANOVA T1/T2 & t test T1/T3**	
Smith et al., 1997	28.0 (4.9)	27.0 (5.4)	36.0 (13.1)	24.0 (8.2)	1.11	t test*	
Sheinkopf & Siegel, 1998	62.8 (27.4)	61.7 (20.2)	89.7 (17.7)	64.3 (25.0)	1.19	Paired t test**	
Smith Groen, & Wynn, 2000	50.5 (11.2)	50.7 (13.9)	66.5 (24.1)	49.7 (19.7)	0.76	t test*	
Eikeseth et al., 2002/ 2007	61.9 (11.3)	65.2 (10.9)	79.1 (18.1) at 12 month FU; 86.9 (SD 25.0) at final FU	69.5 (18.4) at 12 month FU; 71.9 (28.4) at final FU	0.53	t test T1/2** t test T1/3*	
Sallows & Graupner, 2005	50.8 (10.6)	52.1 (8.9)	73.1 (33.1)	79.6 (21.8)	0.23	Not reported but nonsignificant between groups	
Howard et al., 2005	58.5 (18.2)	C1 53.7 (13.5) C2 59.9 (14.8)	89.9 (20.9)	C1 62.1 (19.6) C2 68.1 (15.3) 73 [62-80]	1.28	Multiple regres- sion**	
Cohen et al., 2006 ^c	61.6 (16.4)	59.4 (14.7)	87 [72-98]	73 [62-80]	0.90 ^e	Repeated mea- sures ANCOVA ^{f*}	
Eldevik et al., 2006	41.0 (15.2)	47.2 (14.7)	49.2 (16.6)	44.3 (18.9)	0.27	t test on change scores***	
Magiati et al., 2007 ^c	83.0 (27.9)	65.2 (26.9)	78.4 (17.6)	65.3 (18.0)	0.74	ANCOVA ^c	
Remington et al., 2007	61.4 (16.4)	62.3 (16.6)	73.5 (SD 27.3) at 2 yr FU	58.9 (20.5) at 1 yr FU; 60.1 (SD 27.8) at 2 yr FU	0.48 (at T3)	Repeated mea- sures ANCOVA ^f T1/T2/T3**	

Note. In three studies IQ at 2nd follow-up-Time 3 was also reported. EIBI = early intensive behavioral interventions, C = comparison, C1 = Group 1; C2 = Group 2 in studies with more than comparison group.
^aIQ range given in brackets if SD not reported. ^bBlank indicates no SD or range reported. ^cDifference between EIBI and C IQ controlled for by ANCOVA. ^dNo SD reported for T1 or T2 so not possible to calculate. ^eNo T2 SD reported so T1 SDs used (likely to overestimate effect size at T2). ^fCovarying for Time 1 score.
 p* < .05. *p* < .001.

Table 4. Summary of School Placement at Follow-Up (FU) by Group

Study	Early intensive behavioral intervention (EIBI)	Comparison
Lovaas, 1987; McEachin et al., 1993	At age 7 & at age 11–12 year FUs, 9/19 children enrolled in regular intervention and described as having “normal educational functioning”	At age 7 FU, 1 child and at age 11–12 no child described as “normal educational functioning”
Smith et al., 1997	All children had initial IQ<40; all remained “very much delayed”	All children had initial IQ<40; all remained “very much delayed”
Sheinkopf & Siegel, 1998	5 children reported as attending mainstream, 3 unsupported	No children reported as unsupported in mainstream
Smith, Groen, & Wynn, 2000	4/15 children unsupported in mainstream	No children unsupported in mainstream
Eikeseth et al., 2002, 2007	8/13 children continued to receive EIBI; remainder had full- or part-time support in school, but it is unclear whether this refers to status at final FU	Details unclear
Sallows & Graupner, 2005	8 children unsupported in mainstream but unclear whether these are in clinic-directed EIBI or comparison group	Details unclear
Howard et al., 2005	Not reported	Not reported
Cohen et al., 2006	6/21 children unsupported in mainstream	1/21 children in mainstream (un/supported)
Eldevik et al., 2006	All children remained intellectually impaired (mean IQ<40)	All children remained intellectually impaired (mean IQ<40)
Magiati et al., 2007	23/28 children continuing with EIBI in conjunction with supported placements in mainstream; 5 children in specialist school provision	All children in specialist educational provision
Remington et al., 2007	17/23 children in mainstream for mean of 13 hours per week (not known if un/supported); 5 in special education; 1 in EIBI	10/21 children in mainstream for mean of 22 hours per week (not known if un/supported); 11 in special education for mean of 14 hours per week

the EIBI group) (e.g., Cohen et al., 2006; Sallows & Graupner, 2005), in the majority of studies, the comparison group mean IQ remained the same or decreased slightly.

Second, as noted in almost all reports, although group findings were generally positive in favor of the EIBI groups, there was also considerable individual change, with a minority of children showing marked and significant improve-

ments, and some (see Table 4) achieving *educational independence* (defined as coping without support in mainstream school). For the majority of children, change, although positive, was less dramatic, and some failed to make progress despite, sometimes, thousands of hours of intensive intervention (e.g., Lovaas, 1987). This variability in outcome is also illustrated by the tendency in many studies (see Tables 3 and 5) for variance

Table 5. Summary Data From All Studies on Changes in IQ, VABS, and Language Scores Over Time by Group

Measure ^a	EIBI ^b baseline				EIBI follow-up (FU) ^b				Comparison FU				EIBI change score				Comparison change score				Group change score difference				
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Z ^c	Mann Whitney	
IQ (11)	55.7	13.9	53.6	11.2	74.0	17.6	59.0	15.5	18.3	11.5	-5.0	-31.5	5.4	9.3	-3.3	-27.5									-2.27*
VABS composite SS ^d (8)	61.2	6.4	61.5	6.7	66.2	10.2	56.2	9.3	5.1	6.2	-2.0	-2.0	-2.4	4.5	-10.5	-5.8									-2.58**
Expression SS (7)	49.4	11.9	49.3	7.9	61.1	12.7	50.9	11.9	11.6	10.8	-5	-5.0	1.7	8.5	-7.5	-15.0									-1.79
Comprehension SS (7)	42.8	10.9	45.0	5.3	58.5	10.9	49.6	12.1	15.7	6.0	7.0	-24.0	4.5	12.1	-10.8	-27.0									-1.98*

Note. VABS = Vineland Adaptive Behavior Scales.

^aNumber of studies providing data in parentheses. No differences in group means at baseline. ^bData provided for final FU if more than one FU assessment reported.

^cNonparametric analysis used because change scores highly skewed. ^dStandard score.

* $p < .05$. ** $p < .01$.

scores in both the EIBI and comparison groups to increase over time, indicating greater individual variability over time. As is also evident from Table 5, the range of change scores reported was often similar in both EIBI and comparison groups. Importantly, these data suggest, in addition to any group mean differences, that while some children in both groups do well, others do not. However, only a minority of studies (Cohen et al., 2006; Magiati et al., 2007; Remington et al., 2007; Sallows & Graupner, 2005; Sheinkopf & Siegel, 1998) present clear data on individual patterns of change.

Predictors of Outcome

Individual variability in response to treatment resulted in investigators in 7 out of the 11 studies exploring more systematically variables that seemed to be related to the extent of change (generally as measured by increases in IQ). Some researchers also considered in detail variables that appeared to distinguish between children with the “best” and “poorest” outcomes. Table 6 provides a summary of the intake variables reported as being predictors of favorable outcome, although it is not possible to draw any consistent conclusions from these findings because different variables (or combinations of variables) were assessed in different studies. In 4 studies, initial IQ was strongly positively associated with degree of improvement, but in one (Sallows & Graupner, 2005) initial IQ showed no relationship with outcome. Chronological age was not reported to be associated with outcome, and although the age range of children in these studies was limited (from 3 to 5.5 years), there was nothing in these data to support the claims of some EIBI proponents that in order to be effective, intervention must begin by around 3 years. Initial language level (particularly receptive language) did appear to be associated with outcome in 4 of the 7 studies in which predictive factors were explored. Smith et al. (2000) reported that children with less severe autism symptomatology tend to make better progress, but Remington et al. (2007) found that children with higher levels of behavior problems and autism symptoms at intake were most likely to show the greatest change.

Discussion

Limitations of the Present Analysis

In order to interpret the true effects of intake variables on outcome, reliance on the reports of

Table 6. Predictors of Outcome Following Early Intensive Behavioral Interventions (EIBI)

Study	Predictors of outcome
Lovaas, 1987; McEachin et al., 1993	Initial IQ/MA associated with positive outcome; initial CA did not predict
Smith et al., 1997	None reported; all children had initial IQ < 40
Sheinkopf & Siegel, 1998	Not reported
Smith, Groen, & Wynn, 2000	Children with milder symptoms (PDD ^a diagnosis) tended to do better than those with diagnosis of autism
Eikeseth et al., 2002/2007	At 1 year follow-up (FU), initial and language level predictive; at 3 year FU, initial IQ weakly predictive
Sallows & Graupner, 2005	Initial imitation, language, VABS Daily Living and Socialization Skills, and ADI-R ^b communication best predictors of 3-year outcome. Initial IQ per se did not predict
Howard et al., 2005	Not reported
Cohen et al., 2006 ^c	Not reported
Eldevik et al., 2006	Age at intake not a predictor; initial IQ and language comprehension and expression main predictors
Magiati et al., 2007 ^c	Initial IQ and language comprehension best predictors
Remington et al., 2007	Initial IQ, VABS Composite, Communication, and Social domain scores all positively related outcome; low VABS Motor Scores, high behavior problems, higher autism symptom scores at intake, and fewer hours of intervention in Year 2 also related to improved outcomes.

^aPervasive developmental disorder. ^bAutism Diagnostic Interview-Revised. ^cDerived from Figure 1 in published paper.

group findings is no substitute for a detailed meta-analysis, based on standardized scores for individual children from each of the studies assessed. To date, 11 studies could be identified that met the criteria for inclusion required for the present review and, thus, with cooperation between groups, it is possible that such a meta-analysis could be conducted. Nevertheless, the only variable reported consistently across all studies was IQ, and even this measure was derived from different tests between and within studies.

Other outcome measures varied across studies, and although better standardization of psychometric testing is clearly important, such measures alone do not necessarily provide an adequate global picture of a child's functioning after intervention. There is no a-priori reason why change in IQ should be the principal goal of intervention. Thus, a child might show major increases in IQ over time without improving his or her ability to function in social situations or improving behavior or autistic symptomatology. One likely explanation of this focus on IQ as the primary out-

come variable in the EIBI literature is that the behavioral discrete trial training approach is perhaps best suited to teaching concrete skills as compared to spontaneous social and communication behavior. Few researchers have attempted to assess outcome in terms of behavioral difficulties/severity of autism or the impact of these variables on family life. In contrast, in communication-based interventions, investigators have focused on social and communication outcomes and not IQ (Aldred et al., 2004; Drew et al., 2002; Howlin et al., 2007; Kasari, Freeman, & Paparella, 2006; Yoder & Stone, 2006). Measures of educational status are unreliable indicators of progress because school placement may depend more on parental preferences and/or local educational practices and policies rather than factors inherent in the child. Similarly, in attempting to identify factors related to treatment response, most researchers have focused on easily measurable variables, such as age, initial IQ, or language. The impact of broader variables, such as parental coping ability, family relationships, and stress and support networks, has not

been systematically investigated—and, indeed, will prove a major methodological challenge.

Treatment fidelity is a further issue requiring greater consideration. In only a minority of the 11 studies did researchers provide data on the quality of the alternative therapies offered, and in very few studies was the comparison intervention autism-specific. Thus, aside from not employing the EIBI approach, it is likely that the staff members involved with these children would have less understanding of or expertise in autism and would also be less well-informed about appropriate educational practice for children with autism in general. Details of EIBI interventions are also less detailed than might be expected, for although reassurances are given about the high quality control of these programs, the supporting information provided is generally anecdotal. In community-based studies, perhaps, details of treatment fidelity are less important because these are “effectiveness” trials of implementation in real life. However, in many of the studies reviewed, including the UCLA-based ones, it was even difficult to precisely establish how many hours of intervention children had received. In several investigations, we also found it difficult to establish whether baseline measures had actually been conducted prior to or at the time of the child’s enrollment in the program. Indeed, there were indications in some studies that children had already been exposed to the intervention for some time previously. Thus, Cohen et al. (2006) stated that one inclusion criterion was that children should have received “no more than 400 hours behavioral intervention prior to intake” (p. 146). Moreover, a few investigators (e.g., Magiati et al., 2007; Sallows & Graupner, 2005) mentioned the fact that children in either or both the EIBI or comparison conditions were receiving a number of additional, alternative treatments (e.g., special diets; biological interventions such as Secretin injections or megavitamins; or more controversial therapies such as auditory integration, cranial osteopathy, or pet therapy) as well as being involved in extracurricular activities (e.g., music or play therapy). Few, if any, of these alternative “treatments” have been empirically tested, with the exception of Secretin, where the evidence is clear that it has no empirical basis as an autism treatment (Williams et al., 2005).

A further issue concerns the optimal duration of EIBI programs. Neither the summary overview (Table 5) nor the data provided in the individual studies suggests that length of intervention was

related to outcome. Indeed, Remington et al. (2007) noted that children who responded most positively to intervention had fewer hours of therapy in the second year than those who made the least progress.

In those studies reporting IQ changes at several time points, change scores for the period between the first and subsequent follow-up were generally much smaller than the changes between baseline and Follow-Up 1 (see Table 7 for a summary of studies with three or more time points). For example, in Lovaas’ (1987) study, the mean increase in IQ between initial assessment and follow-up at 6 to 7 years was 30 points; at the next follow-up, when children were 11 to 12 years of age, the further increase in IQ was 1.5 points. Eikeseth et al. (2002) found that the first year increase was around 17 points, with an 8-point increase in the following year. Cohen et al. (2006) reported that mean IQ rose approximately 16 points in the first year but then increased by about 3 to 4 points the following year and about 5 points in the final year (data in this study are presented graphically and, thus, not easy to read precisely). In Remington et al.’s (2007) study, there was a mean increase in IQ of 8 points in the first year and a further increase of 4 points in the subsequent year. Although these subsequent increases are not negligible, they do suggest that the main impact of intervention is in the first year, and, thereafter, increases, at least in IQ, tend to plateau. Of course, IQ is not the only variable of interest and, as noted above, is not necessarily the most important indicator of treatment effectiveness. However, in those 3 studies that also showed changes in VABS scores over several time points, the diminution in progress after the first year was even more marked, with few changes in standard scores occurring after 12 months (Eikeseth et al., 2002, 2007; Remington et al., 2007; Sallows & Graupner, 2005). Cohen et al. (2006) reported a steady increase in expressive language from 12 to 24 months post-intake, but this progressive increase was not replicated in the only other study (Remington et al., 2007) in which language scores over more than two time points were tracked.

Implications for Future Intervention Research

The need to improve the quality of autism intervention research generally has been stressed in a number of recent reviews, with specific recommendations concerning the need to address issues such as small sample sizes, nonrandomized

Table 7. IQ Changes in Early Intensive Behavioral Interventions (EIBI) Group for Studies Presenting Data Over Three or More Time Points

Study	Time 1		Time 2		Time 3		Time 4	
	Average age mean ^a	IQ (SD)	Average (approx.) age mean	IQ (SD)	Average (approx.) age mean	IQ (SD)	Average (approx.) age mean	IQ
Lovaas, 1987;								
McEachin et al., 1993	35	53 ^b	6–7 years	83.3 ^b	11.5 years	84.5 (32.4)	N/A ^c	
Eikeseth et al., 2002/2007	66	61.9 (11.3)	79 months	79.1 (18.1)	98 months	86.9 (25.0)	N/A	
Cohen et al., 2006	30 ^b	61.6 (16.4)	Initial age + 1 year	77 ^c	Initial age + 2 years		Initial age + 3 years	87 ^d
Remington et al., 2007	36	61.4 (16.4)	Initial age + 1 year	68.8 (20.5)	Initial age + 2 years	73.5 (27.3)	N/A	

^aIn months. ^bSD not provided; age at diagnosis. ^cNot applicable (only three time points reported). ^dDerived from Figure 1 in published paper.

assignment, inadequate information on the characteristics of participants and therapy, and the limitations of outcome measures (see Charman et al., 2003; Lord et al., 2005; Rogers, 1998). The very small number of evaluations conducted by researchers who had no direct involvement in the therapy is another source of concern. Although the research base for studies involving behavioral interventions is, in general, methodologically stronger than for many other types of intervention, it is clear from our review that none of even the better controlled EIBI studies met the most rigorous standards for intervention methodology (National Research Council, 2001). Moreover, differences in therapeutic intensity, in participants, and in intake and outcome measures have led to considerable disagreement and controversy among researchers and professionals. In turn, this has resulted in confusion and distress for many parents of young children with autism spectrum disorders, who are faced with conflicting claims of effectiveness when seeking appropriate treatment.

There is little question now that early intensive behavioral intervention is highly effective for some children. However, gains are not universal, and some children make only modest progress while others show little or no change, sometimes after extremely lengthy periods in treatment. Guralnick (2005) noted that one of the most important tasks in autism early intervention research is to understand why outcomes vary so dramatically across different children, and the crucial question to be systematically addressed is for which children is EIBI most and least effective? Unfortunately, because of the focus on group differences, existing research provides only limited information on the outcome for individual children and few data on moderators or mediators of therapy (Lord et al., 2005; Yoder & Compton, 2004). It is almost impossible to draw reliable conclusions about possible child, family, or environmental variables associated with outcome when most studies involve relatively small numbers of participants. More detailed, multisite studies with large samples (see the current United Kingdom Pre-School Autism Communication Trial (2008) for a rare exception) and the sharing of data on individual cases are necessary for such investigations. At the very least, when sample size is small, as is currently the case for most treatment evaluation studies, it is important that more detailed information on individual differences is presented, in order to ensure that group differences are not due

mainly to improvements in a small subgroup of individuals.

Furthermore, from the papers reviewed here, there is an indication that the immediate impact of EIBI reduces over time, with the first year of intervention appearing to produce the most substantial gains, at least in IQ. EIBI is a relatively expensive form of intervention, and the cost can be a major deterrent for many parents and for education authorities who may be asked to provide funding. The demands on family life are also considerable. Clearer evidence concerning both the optimal duration of therapy and the age at which it should begin could result in the development of better targeted, more cost effective programs that could then be made more widely accessible to families. There is also a need to demonstrate that EIBI is substantially more effective than alternative, high quality autism specific interventions, such as specialized preschool provision. In the majority of the studies included in this review, the alternative intervention has generally been of lower intensity and/or lesser quality (certainly in terms of autism expertise) than the EIBI program to which it is compared. In addition, the failure to control for time in intervention means that in some studies the fact that EIBI appears to be more effective than the comparison condition could be due simply to differences in treatment intensity, not quality. As noted above, a systematic meta-analysis is required in order to explore both child and intervention variables that may be related to outcome. However, given the variability of data, it is likely that such an analysis at the present time could only assess the role of a very limited number of variables.

Conclusions and Recommendations for Future Research

Since we initially submitted this paper, one other comprehensive review of early intensive behavioral interventions based on the UCLA Young Autism Project model has been published (Reichow & Wolery, 2008). Although that analysis was focused on some different issues and includes papers that did not meet our criteria for inclusion, the principal conclusion parallels our own; that is, there is strong evidence that EIBI is effective for some, but not all, children with autism spectrum disorders, and there is wide variability in response to treatment. The data we present in this review also suggest that if gains are made, they are likely

to be greatest in the first 12 months of intervention. It is evident, too, that in recent studies researchers have followed some of the recommendations arising from earlier reviews (e.g., Lord et al., 2005; Rogers, 1988; Smith et al., 2007), in particular by including a wider range of child measures, especially those related to social and communication skills. However, broader measures of family functioning are rarely included and as yet there are no data on how these may affect outcome. There remains a dearth of randomized control trials, which are needed in order to provide unbiased evidence of efficacy.

If research and practice in the area of early intervention for children with autism are to improve, it is clear that increasing the number of randomized control trials should be made a high priority. Nevertheless, because the number of primary outcome measures of such trials is, of necessity, limited, other case control comparison studies will continue to be needed—but with adequate sample sizes in order to explore individual variables associated with outcome. For such studies to be informative, the following minimum requirements need to be adequately addressed: (a) Baseline data on participants should be collected immediately prior to, or at the very beginning, of the intervention program. (b) The age at which individual children begin treatment, the duration and intensity of treatment (measured in hours per week), and the exact time to follow-up, should be made explicit. (c) The same data should be available for children in the comparison conditions, with as much additional information as possible provided in order to enable readers to assess the true quality of that intervention. (d) Data on child variables need to be better standardized. Assessments used should include those currently in general use (IQ, VABS, language comprehension and expression), but additional measures of behavioral disturbance (e.g., the Developmental Behaviour Checklist: Einfeld & Tonge, 1995) are also required. Standard scores, age equivalent scores, and raw scores (at least in an online appendix) should be reported for all assessments. (e) Diagnostic status should be clearly defined using the ADI-R and ADOS-G because these instruments can also be used to monitor changes in autistic symptomatology over time. (f) Measures of family functioning need to be included; not only data on marital and economic status, but also measures (such as those employed by Remington et al., 2007) of parental well-being and support.

Finally, it is important to recognize that autism is a highly complex and heterogeneous disorder characterized by impairments in social behavior, in communication, and in many aspects of learning, together with fundamental problems in acquiring functional, adaptive, and flexible behaviors. Although behavioral approaches are an important element of any comprehensive program (perhaps especially in the early years), other elements that focus more specifically on social development and communication will also be required for optimal effectiveness. This will also require more development of appropriate measures of social and communicative competence to assess changes in these skills (Lord et al., 2005). Moreover, wide variation in child characteristics and in responsiveness to treatment may require a far more individualized approach to intervention than can be delivered using manualized treatment packages.

There is good evidence now, from a number of randomized control trials (Aldred et al., 2004; Howlin et al., 2007; Kasari et al., 2006; Yoder & Stone, 2006), that other, nonintensive interventions, particularly those with a focus on communication and joint social interaction, can have a significant and positive impact on children's functioning. A switch of focus to examining the comparative effects of interventions with a strong evidence base, called equivalence trials, rather than simply comparing high quality interventions with low quality/low intensity alternatives, is likely to be of benefit to many more children and their families. Assessing what treatments work for which children and identifying the individual characteristics that predict responsiveness to specific programs and approaches are the challenges that lie ahead.

References (* indicates reference is highlighted in tables)

- Aldred, C., Green, J., & Adams, C. (2004). A new social communication intervention for children with autism: Pilot randomized controlled treatment study suggesting effectiveness. *Journal of Child Psychology and Psychiatry*, 45, 1420–1430.
- Anderson, S. R., & Romanczyk, R. G. (1999). Early intervention for young children with autism: Continuum-based behavioral models. *Journal of the Association for Persons with Severe Handicaps*, 24, 162–173.
- Baer, D. (1993). Quasi random assignment can be as convincing as random assignment. *American Journal on Mental Retardation*, 97, 373–375.
- Barbera, M. L., & Rasmussen, T. (2007). *The verbal behavior approach: How to teach children with autism and related disorders*. London: Jessica Kingsley.
- Bondy, A. S., & Frost, L. A. (1998). The Picture Exchange Communication System. *Seminars in Speech and Language*, 19, 373–389.
- Charman, T. (2004). Matching preschool children with autism spectrum disorders and comparison children for language ability: Methodological challenges. *Journal of Autism and Developmental Disorders*, 34, 59–64.
- Charman, T., & Howlin, P. (with Aldred, C., Baird, G., degli Espinosa, F., Diggle, T., Kovshoff, H., Law, J., et al.). (2003). Research into early intervention for children with autism and related disorders: Methodological and design issues. *Autism*, 7, 217–225.
- *Cohen, H., Amerine-Dickens, M. S., & Smith, T. (2006). Early Intensive Behavioral Treatment: Replication of the UCLA model in a community setting. *Journal of Developmental and Behavioral Pediatrics*, 27, S2, 145–155.
- Drew, A., Baird, G., Baron-Cohen, S., Cox, A., Slonims, V., Wheelwright, S., Swettenham, J., Berry, B., & Charman, T. (2002). A pilot randomised control trial of a parent training intervention for pre-school children with autism—Preliminary findings and methodological challenges. *European Child and Adolescent Psychiatry*, 11, 266–272.
- Dunlap, G., Kern-Dunlap, L., Clark, S., & Robbins, F. (1991). Functional assessment, curricular revision, and severe problem behaviors. *Journal of Applied Behavior Analysis*, 4, 387–397.
- Einfeld, S. L., & Tonge, B. J. (1995). The Developmental Behaviour Checklist: The development and validation of an instrument to assess behavioural and emotional disturbance in children and adolescents with mental retardation. *Journal of Autism and Developmental Disorders*, 25, 81–104.
- *Eikeseth, S., Smith, T., Jahr, E., & Eldevik, S. (2002). Intensive behavioral treatment at school for 4- to 7-year-old children with autism: A one-year comparison controlled study. *Behavior Modification*, 26, 49–68.
- *Eikeseth, S., Smith, T., Jahr, E. & Eldevik, S.

Early intensive behavioral interventions

P. Howlin, I. Magiati, and T. Charman

- (2007). Outcome for children who began intensive behavioral treatment between ages 4–7: A comparison controlled study. *Behavior Modification, 31*, 264–278.
- *Eldevik, S., Eikeseth, S., Jahr, E., & Smith, T. (2006). Effects of low intensity behavioral treatment for children with autism and mental retardation. *Journal of Autism and Developmental Disorders, 36*, 211–224.
- Greenspan, S. L., & Wieder, S. (2003). *Engaging autism: The Floortime Approach to helping children relate, communicate and think*. Jackson, TN: Perseus Books.
- Guralnick, M. (2005). Early intervention for children with intellectual disabilities: Current knowledge and future prospects. *Journal of Applied Research in Intellectual Disabilities, 18*, 313–324.
- Handleman, J. S., Harris, S. L., Arnold, M., Gordon, R., & Cohen, M. (2006). The Douglass Developmental Disabilities Center: Options for school-age students. In J. S. Handleman, & S. L. Harris (Eds.), *School-age education programs for children with autism. The in-between years* (pp. 89–113). Austin, TX: Pro-Ed.
- *Howard, J. S., Sparkman, C. R., Cohen, H. G., Green, G., & Stanislaw, H. (2005). A comparison of intensive behavior analytic and eclectic treatments for young children with autism. *Research in Developmental Disabilities, 26*, 359–383.
- Howlin, P. (2006). Augmentative and alternative communication systems for children with autism. In T. Charman & W. Stone (Eds.), *Social and communication development in autism spectrum disorders* (pp. 236–266). New York: Guilford Press.
- Howlin, P., Gordon, K., Pasco, G., Wade, A., & Charman, T. (2007). A group randomised, controlled trial of the Picture Exchange Communication System for children with autism. *Journal of Child Psychology and Psychiatry, 48*, 473–481.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12–19.
- Kasari, C., Freeman, S., & Paparella, T. (2006). Joint attention and symbolic play in young children with autism: A randomized controlled intervention study. *Journal of Child Psychology and Psychiatry, 47*, 611–620.
- Kaufman, B. N. (1994). *Son-Rise: The miracle continues*. Tiburon, CA: Kramer.
- Levy, S., Kim, A.-H., & Olive, M. (2006). Interventions for young children with autism: A synthesis of the literature. *Focus on Autism and Other Developmental Disabilities, 21*, 55–62.
- Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., Pickles, A., & Rutter, M. (2000). The Autism Diagnostic Observation Schedule-Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30*, 205–223.
- Lord, C., Rutter, M., & Le Couteur, A. (1994). Autistic Diagnostic Interview–Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*, 659–686.
- Lord, C., Wagner, A., Rogers, S., Szatmari, P., Aman, M., Charman, T., et al. (2005). Challenges in evaluating psychosocial interventions for autistic spectrum disorders. *Journal of Autism and Developmental Disorders, 35*, 695–711.
- *Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology, 55*, 3–9.
- Lovaas, O. I., Ackerman, A., Alexander, D., Firestone, P., Perkins, M., Young, D. B., et al. (1981). *Teaching developmentally disabled children: The ME book*. Austin, TX: Pro-Ed.
- *Magiati, I., Charman, T., & Howlin, P. (2007). A two-year prospective follow-up study of community-based early intensive behavioural intervention and specialist nursery provision for children with autism spectrum disorders. *Journal of Child Psychology and Psychiatry, 48*, 803–812.
- Magiati, I., & Howlin, P. (2001). Monitoring the progress of pre-school children with autism enrolled in early intervention programmes: Problems of cognitive assessment. *Autism, 5*, 399–406.
- Maine Administrators of Services for Children With Disabilities (MADSEC). (1999). Report of the MADSEC Autism Task Force, Manchester, ME: Author.
- Maurice, C., Green, G., & Foxx, R. M. (Eds.), (2001). *Making a difference: Behavioral intervention for autism*. Austin, TX: Pro-Ed.

- Maurice, C., Green, G., & Luce, S. (Eds.). (1996). *Behavioral intervention for young children with autism: A manual for parents and professionals*. Austin, TX: Pro-Ed.
- *McEachin, J. J., Smith, T., & Lovaas, O. I. (1993). Long-term outcome for children with autism who received early intensive behavioral treatment. *American Journal on Mental Retardation*, 97, 359–372
- Mesibov, G. B. (1993). Treatment outcome is encouraging: Comments on McEachin et al. *American Journal on Mental Retardation*, 97, 379–380.
- Mundy, P. (1993). Normal versus high-functioning status in children with autism (Comments on McEachin, Smith, and Lovaas). *American Journal on Mental Retardation*, 97, 373–391.
- National Initiative for Autism: Screening and Assessment (NIASA). (2003). *National Autism Plan for Children (NAPC). Plan for the identification, assessment, diagnosis and access to early interventions for pre-school and primary school aged children with ASD*. London: National Autistic Society.
- National Research Council. (2001). *Educating children with autism* (Committee on Educational Interventions for Children With Autism, Division of Behavioral and Social Sciences and Education). Washington, DC: National Academy Press.
- New York State Department of Health Early Intervention Program. (2000). *Clinical practice guideline: Report of the Recommendations for Autism/Pervasive Developmental Disorders: Assessment and intervention for young children (age 0–3 years)*. Retrieved October, 10, 2000, from <http://www.health.state.ny.us/nysdoh/eip/autism/index.htm>
- New Zealand Ministry of Health. (2006). *Evidence based guidelines for autism spectrum disorders* (Draft report). Wellington, NZ: Author.
- Partington, J. W., & Sundberg, M. L. (1998). *The Assessment of Basic Language and Learning Skills: An assessment, curriculum guide and tracking system for children with autism or other developmental disabilities*. Danville, CA: Behavior Analysts.
- Pre-School Autism Communication Trial (PACT). (2008). Retrieved September 29, 2008, from <http://www.medicine.manchester.ac.uk/pact>
- Quill, K., Gurry, S., & Larkin, A. (1989). Daily life therapy: A Japanese model for educating children with autism. *Journal of Autism and Developmental Disorders*, 19, 625–635.
- Reichow, B., & Wolery, M. (2008). Comprehensive synthesis of early intervention behavioral interventions for young children with autism based on the UCLA Young Autism Project model. *Journal of Autism and Developmental Disorders*. DOI 0. 1007/S10803-008-0596-0
- *Remington, B., Hastings, R. P., Kovshoff, H., degli Espinosa, F., Jahr, E., Brown, T., et al. (2007). Early intensive behavioral intervention: Outcomes for children with autism and their parents after two years. *American Journal on Mental Retardation*, 112, 418–438.
- Research Autism. (2007). Retrieved August 9, 2007, from <http://www.researchautism.net/pages/welcome/home.ikml>
- Roberts, J. M. (2004). *A review of the research to identify the most effective models of best practice in the management of children with autism spectrum disorders*. Centre for Disability Studies, University of Sydney, Australia.
- Rogers, S. J. (1998). Empirically supported comprehensive treatments for young children with autism. *Journal of Clinical Child Psychology*, 27, 168–179.
- Rogers, S. J., Hayden, D., Hepburn, S., Charlifue-Smith, R., Hall, T., & Hayes, A. (2006). Teaching young nonverbal children with autism useful speech: A pilot study of the Denver model and PROMPT interventions. *Journal of Autism and Developmental Disorders*, 36, 1007–1024.
- *Sallows, G. O., & Graupner, T. D. (2005). Intensive behavioral treatment for children with autism: Four-year outcome and predictors. *American Journal on Mental Retardation*, 110, 417–438.
- Schopler, E. (1997). Implementation of TEACCH philosophy. In D. J. Cohen & F. R. Volkmar (Eds.), *Handbook of autism and pervasive developmental disorders*. (2nd ed., pp. 767–798). New York: Wiley.
- Schreibman, L., & Koegel, R. L. (2005). Training for parents of children with autism: Pivotal responses, generalization, and individualization of interventions. In E. D. Hibbs & P. S. Jensen (Eds.), *Psychosocial treatment for child and adolescent disorders: Empirically based strategies for clinical practice* (2nd ed., pp. 605–631). Washington, DC: American Psychological Association.
- Scottish Intercollegiate Guidelines Network

- (SIGN). (2007). *Assessment, diagnosis, and clinical interventions for children and young people with autism spectrum disorders*. Edinburgh Scotland: National Health Service Education for Scotland.
- Shea, V. (2004). A perspective on the research literature related to early intensive behavioural intervention (Lovaas) for young children with autism. *Autism, 8*, 349–367.
- *Sheinkopf, S. J., & Siegel, B. (1998). Home-based behavioral treatment of young children with autism. *Journal of Autism and Developmental Disorders, 28*, 15–23.
- Shields, J. (2001). The NAS Early Bird programme: Partnership with parents in early intervention. *Autism, 5*, 49–56.
- Skinner, B. F. (1953). *Science and human behavior*. New York: Macmillan.
- Smith, T., Donahoe, P. A., & Davis, B. J. (2006). The UCLA Young Autism Project. In J. S. Handleman & S. L. Harris (Eds.), *Preschool education programs for children with autism* (2nd ed., pp. 29–48). Austin, TX: Pro-Ed.
- Smith, T., Eikeseth, S., Klevstrand, M., & Lovaas, I. (1997). Intensive behavioral treatment for preschoolers with severe mental retardation and pervasive developmental disorder. *American Journal on Mental Retardation, 102*, 238–249.
- *Smith, T., Groen, A. D., & Wynn, J. W. (2000). Randomized trial of intensive early intervention for children with pervasive developmental disorder. *American Journal on Mental Retardation, 105*, 269–285.
- *Smith, T., Scahill, L., Dawson, G., Guthrie, D., Lord, C., Odom, S., et al. (2007). Designing research studies on psychosocial interventions in autism. *Journal of Autism and Developmental Disorders, 37*, 354–366.
- Sparrow, S. S., Balla, D., & Cicchetti, D. V. (1984). *Vineland Adaptive Behavior Scales*. Circle Pines, MN: American Guidance Service.
- Sundberg, M. L., & Partington, J. W. (1999). The need for both discrete trial and natural environmental language training for children with autism. In P. M. Ghezzi, W. L. Williams, & J. E. Carr (Eds.), *Autism behavior analytic perspectives* (pp. 139–156). Reno, NV: Context Press.
- Sussman, F. (1999). *More than words: Helping parents promote communication and social skills in children with autism spectrum disorders*. Toronto: Hanen Center.
- Wheeler, J. J., Baggett, B. A., Fox, J., & Blevins, L. (2006). Treatment integrity: A review of intervention studies conducted with children with autism. *Focus on Autism and Other Developmental Disabilities, 21*, 45–54.
- Williams, K. W., Wray, J. J., & Wheeler, D. M. (2005). Intravenous secretin for autism spectrum disorder. *Cochrane Database Systematic Reviews, 20*, CD003495.
- Wirt, R. D., Lachar, D., Klinedinst, J. K., & Seat, P. D. (1984). *Multi-dimensional description of child personality: A manual for the Personality Inventory for Children*. Los Angeles: Western Psychological Services.
- Yoder, P. J., & Compton, D. (2004). Identifying predictors of treatment response. *Mental Retardation and Developmental Disabilities Research Reviews, 10*, 162–168.
- Yoder, P. J., & Stone, W. L. (2006). Randomized comparison of two communication interventions for preschoolers with autism spectrum disorders. *Journal of Consulting and Clinical Psychology, 74*, 426–235.

Received 9/3/07, accepted 1/16/08.

Editor-in-charge: William E. MacLean, Jr.

Correspondence regarding this article should be sent to Patricia Howlin, Department of Psychology, Institute of Psychiatry, Henry Wellcome Bldg., PO 77, De Crespigny Park, London SE5 8AF, UK. E-mail: patricia.howlin@iop.kcl.ac.uk