Research in Autism Spectrum Disorders 6 (2012) 673–682

Contents lists available at SciVerse ScienceDirect

Research in Autism Spectrum Disorders

Journal homepage: http://ees.elsevier.com/RASD/default.asp

Effectiveness of large-scale community-based Intensive Behavioral Intervention: A waitlist comparison study exploring outcomes and predictors

Helen E. Flanagan [a,*], Adrienne Perry [a], Nancy L. Freeman [b]

[a] York University, Department of Psychology, Behavioural Sciences Building, 4700 Keele Street, Toronto, ON, Canada M3J 1P3
[b] Surrey Place Centre, 2 Surrey Place, Toronto, ON, Canada M5S 2C2

ARTICLE INFO

Article history:
Received 10 September 2010
Received in revised form 21 September 2011
Accepted 21 September 2011

Keywords:
Early Intensive Behavioral Intervention
Autism
Children
Effectiveness study
Treatment
Predictors
Outcomes

ABSTRACT

File review data were used to explore the impact of a large-scale publicly funded Intensive Behavioral Intervention (IBI) program for young children with autism. Outcomes were compared for 61 children who received IBI and 61 individually matched children from a waitlist comparison group. In addition, predictors of better cognitive outcomes were explored ($n = 142$). Although random assignment did not take place, a standardized waitlist management system was used that did not include any prioritization other than time of referral. Groups did not differ significantly on available measures at intake. The treatment period tended to be longer than the waitlist period and this difference was controlled in analyses. At exit, the IBI group had better outcomes in all measured areas, with milder autism severity, higher adaptive functioning, and higher cognitive skills. Younger initial age predicted better cognitive outcomes in the IBI group but not the Waitlist group. Higher initial adaptive skills predicted better outcomes similarly in the two groups. Results support the effectiveness of community-based IBI and suggest that earlier age at treatment onset may increase the likelihood of better outcomes relative to comparison conditions.

1. Introduction

1.1. The impact of Intensive Behavioral Intervention

For many years, behavioral interventions have shown promise as an effective approach to increasing the skills of people with autism (e.g., Ferster & DeMyer, 1961; Lovaas, 1967; Lovaas, Koegel, Simmons, & Long, 1973). In 1987, Lovaas demonstrated that young children with autism could make remarkable gains after receiving 40 h/week of early Intensive Behavioral Intervention (IBI) for at least 2 years (Lovaas, 1987). Since this time, a growing body of research has supported IBI relative to comparison conditions such as low intensity behavioral intervention (e.g., Smith, Groen, & Wynn, 2000), special education (e.g., Howard, Sparkman, Cohen, Green, & Stanislaw, 2005), and intensive eclectic intervention (e.g., Eikeseth, Smith, Jahr, & Eldevik, 2007; Howard et al., 2005). Meta-analyses provide varying interpretations of treatment effects (Eldevik et al., 2009; Howlin, Magiati, & Charman, 2009; Makrygianni & Reed, 2010; Reichow & Wolery, 2009). In one recent meta-analysis that carefully examined the impact of IBI relative to comparison conditions, IBI was found to be associated

Abbreviations: IBI, Intensive Behavioral Intervention; TPAS, Toronto Partnership for Autism Services; CARS, Childhood Autism Rating Scale; VABS, Vineland Adaptive Behavior Scales.

* Corresponding author. Current address: IWK Health Centre, 5850/5980 University Avenue, Halifax, NS, Canada B3K 6R8. Tel.: +1 902 470 7732.
E-mail addresses: helen.flanagan@iwk.nshealth.ca, hflanagan1@gmail.com (H.E. Flanagan).

1750-9467/$ – see front matter
doi:10.1016/j.rasd.2011.09.011

with large mean gains in cognitive skills and moderate mean gains in adaptive functioning (Reichow & Wolery, 2009). Comprehensive behavioral treatments are considered to be an established intervention for children with autism (National Autism Center, 2009), and expert consensus panels have recommended IBI as the treatment of choice for this population (e.g., New York State Department of Health, 1999).

Although IBI has a strong body of empirical support, more research is needed to better understand its effectiveness in community settings. Most existing comparison group studies have examined the impact of IBI under carefully regulated conditions. Treatment has typically been supervised by researchers or supported by University centers (e.g., Howard et al., 2005; Lovaas, 1987) and children have often been excluded if they are over 3 or 4 years of age at intake (e.g., <48 months in Cohen, Amerine-Dickens, & Smith, 2006; Howard et al., 2005; <42 months in Sallows & Graupner, 2005) or have severe cognitive delays (e.g., ratio IQ < 35 in Cohen et al., 2006; Sallows & Graupner, 2005; Smith et al., 2000). The impact of IBI may differ when it is provided under less stringent conditions and with more heterogeneous samples. For example, two studies suggest that self-funded parent-managed IBI may produce relatively small gains (Bibby, Eikeseth, Martin, Mudford, & Reeves, 2002; Magiati, Charman, & Howlin, 2007). In a study examining the effectiveness of IBI provided by one of a number of service providers in the United Kingdom, children who received IBI had better cognitive outcomes (relatively large effect size), higher daily living skills, and better performance on a number of language and social measures than a group who received standard community and education services. However, overall adaptive functioning and parent-rated autism symptoms did not differ across groups (Remington et al., 2007).

Additional research is also needed to better understand variables that might predict which children are most likely to succeed in IBI programs. In previous research, higher initial IQ has been linked to better cognitive outcomes (Bibby et al., 2002; Eikeseth et al., 2007; Harris & Handleman, 2000; Lovaas, 1987; Sallows & Graupner, 2005), less restrictive school placements (Harris & Handleman, 2000; Lovaas, 1987), and higher adaptive functioning (Eikeseth et al., 2007). Similarly, higher initial adaptive scores have been linked to better language/social (Sallows & Graupner, 2005) and cognitive (Remington et al., 2007) outcomes. Younger age at treatment intake may also predict higher IQ (Harris & Handleman, 2000) and higher adaptive functioning (Goin-Kochel, Myers, Hendricks, Carr, & Wiley, 2007), although this has not always been found (Eikeseth et al., 2007; Lovaas, 1987). Similarly, milder initial autism severity has been linked to better outcomes in some (Ben-Itzchak & Zachor, 2007; Sallows & Graupner, 2005), but not all (Harris & Handleman, 2000; Remington et al., 2007) previous research. Other potentially important predictor variables include faster early learning rate during IBI (Goin-Kochel et al., 2007; Sallows & Graupner, 2005) and the model of training of supervisors (Reichow & Wolery, 2009).

In previous studies exploring predictor variables, sample sizes have been relatively small and researchers have not always controlled for potentially intercorrelated variables (e.g., Harris & Handleman, 2000; Smith et al., 2000). Although a few studies have examined predictors of outcome within both IBI and comparison groups (e.g., Eikeseth et al., 2007; Lovaas, 1987), none has examined interactions between predictors and group membership. The latter approach has been recognized as an ideal strategy to determine predictors of treatment response (Yoder & Compton, 2004).

1.2. The Ontario IBI initiative and the present study

This study examined the effectiveness of a large-scale community-based IBI program, the Toronto Partnership for Autism Services (TPAS), using a waitlist comparison group. TPAS is one of nine regional programs that comprise the Ontario IBI Initiative, a province-wide IBI program in Canada that was launched in 2000 (see Perry, 2002; Perry et al., 2008). Initial research supports the effectiveness of the Ontario IBI Initiative. For example, Freeman and Perry (2010) reported significant improvements in cognitive skills and significant reductions in autism symptoms in a sample of 89 children receiving intervention through TPAS. Perry et al. (2008) reported similar findings in a larger sample of 332 children from all nine regional programs throughout the province. Over the course of IBI, there were significant reductions in autism severity, small improvements in standard scores for some domains of adaptive functioning, and significant improvements in cognitive skills. In a second paper based on the same sample, higher initial skill level and younger initial age were found to be predictors of better outcome (Perry et al., 2011).

The current study builds on the above research by examining the impact of the Ontario IBI Initiative using a waitlist comparison group. First, children who received IBI through TPAS were individually matched to children on the waitlist, and outcome measures assessing cognitive skills, adaptive skills and autism severity were compared (see Matson, 2007). Secondly, hierarchical multiple regression was used to examine predictors of better cognitive outcomes. Regression analyses focused on the role of four possible predictors: duration of treatment, initial age, initial adaptive skills, and initial autism severity. Interactions of each variable with group membership were examined to determine whether predictor variables had a different relationship with outcomes for the two intervention conditions. Unfortunately, we could not include initial cognitive skill level as a predictor, as data about initial cognitive functioning were not available.

2. Materials and methods

2.1. Procedure

Information from most measures was obtained through file review (ethical approval was obtained stating that file data could be used for research purposes). All children had either received IBI at TPAS (referred to as the IBI group) or been on a waitlist for intervention at TPAS (referred to as the Waitlist group). At TPAS, all children are placed on a waitlist before they begin IBI due to a high demand for service. Transitions from the waitlist to treatment are based strictly on date of referral, and are not influenced by other child and family variables (as per program guidelines, Ministry of Children and Youth Services, 2006).

File review data were acquired regarding children's skills prior to the beginning of the treatment or waitlist period (referred to as Time 1) and following time in treatment or on the waitlist (referred to as Time 2). Parents in the Waitlist group were also contacted by telephone to obtain basic information about demographics and community services received. Unfortunately, we were unable to obtain ethical approval to contact families in the IBI group to obtain demographic and service data. Many of these families were no longer clients of TPAS, and their contact information could not be released.

2.2. Sample characteristics

A dataset was created using information in the clinical files of children who had completed IBI or left the waitlist at TPAS within the previous 4 years ($n = 196$). Children were removed from this dataset if they did not meet the following inclusion criteria:

(a) In IBI or on the waitlist for at least 12 months (excluded 8 children).
(b) Complete information available about adaptive functioning (Time 1 and Time 2), autism severity (Time 1 and Time 2) and cognitive skills (Time 2), with all measures at the same time point completed within 3 months of one another (excluded 37 children).
(c) If on the waitlist: received fewer than 10 h/week of IBI from private agencies (excluded 4 children).
(d) If received IBI: received IBI for at least 80% of the interval between Time 1 and Time 2 testing (excluded 4 children who had received initial testing long before their IBI program began).

One additional child in the Waitlist group was also excluded because he was much older than all other children at intake. This procedure resulted in a sample of 142 children (79 who had received IBI and 63 who had been on the waitlist). As a group, these 142 children did not differ significantly from children excluded from analyses ($p > .05$ for comparisons regarding initial age, initial autism severity, and initial adaptive level). Some of the children in the IBI group were also included in previous studies exploring the impact of the Ontario IBI Initiative (16% were included in Freeman & Perry, 2010; 27% were included in Perry et al., 2008, 2011).

The Ontario IBI Initiative accepts children of varied ages and skill levels, and does not exclude children with comorbid diagnoses or require parental involvement. Prior to referral to TPAS, each child was diagnosed with an autism spectrum disorder by a qualified professional in the community. At TPAS, children were re-assessed by an experienced psychologist with expertise in autism. Diagnoses were obtained using the Childhood Autism Rating Scale (CARS), in combination with clinical observation and a diagnostic and adaptive interview. All children met diagnostic criteria for an autism spectrum disorder towards the severe end of the autism spectrum (50% had autism and 50% had PDD-NOS). Most children in this sample began IBI or the Waitlist with significant delays. Only 6% had adaptive standard scores of 70 or higher (the highest was 77), and only 44% had scores of 55 or higher.

Toronto is a culturally diverse city, and TPAS serves children from a wide range of national, economic, and linguistic backgrounds. Demographic information was obtained for 87% of children in the Waitlist group, but was not available for the IBI group. Children in both groups were drawn from the same sample, and there is no reason to believe that groups would differ significantly with respect to demographic background. Based on available data, parents of most children were born outside of Canada (84% had parents born in one of 21 other countries). Parents typically spoke either English (44%) or English and another language (42%) at home, and were typically married (67%). Parental education level varied, with neither parent attending college or university in 29% of cases, at least one parent attending college or university in 51% of cases, and at least one parent completing a professional or graduate degree in 20% of cases.

2.3. Measures

Assessments were carried out based on best practice guidelines for children with autism (Filipek et al., 1999). All testing was conducted by Master's level psychometrists or graduate-level psychology students working under the supervision of registered psychologists. As data were collected for clinical reasons, assessors knew whether or not children had received IBI.

2.3.1. Autism severity

Autism severity was measured at both Time 1 and Time 2 using the Childhood Autism Rating Scale (CARS; Schopler, Reichler, & Renner, 1988). The CARS is a behavior observation measure with established reliability and validity (Eaves & Milner, 1993; Perry, Condillac, Freeman, Dunn Geier, & Belair, 2005) that provides a total score falling within either the non-autism (<30), mild-moderate autism (30–36.5), or severe-autism (≥ 37) range. Prior to administering the CARS, assessors were trained to reliability (procedure was similar to that described in Perry et al., 2005).

2.3.2. Adaptive skills

At both Time 1 and Time 2, adaptive behavior was assessed using the Vineland Adaptive Behavior Scales (VABS; Sparrow, Balla, & Cicchetti, 1984; for two children the Vineland-II [Sparrow, Cicchetti, & Balla, 2005] was used). The VABS is a psychometrically valid semi-structured interview with items in four domains that assess communication, socialization, daily living skills, and motor skills. The motor skills domain is only applicable for children under 6 years of age. To ensure consistency when comparing children of various ages, motor skills were not included in the current study.

Most previous studies have explored changes in adaptive functioning using standard scores (e.g., Cohen et al., 2006; Sallows & Graupner, 2005). However, there are limitations associated with using standard scores in samples that include younger, lower functioning children, such as basal effects and decreases in the floor of the scale as children age (Carter et al., 1998; Perry, Flanagan, Dunn Geier, & Freeman, 2009). For example, the floor of the communication domain on the VABS decreases from a score of 53 at 24 months of age to a score of 24 at 84 months of age. These effects can create strong negative correlations between age and adaptive standard scores (e.g., r (140) = −.79, $p < .001$ for composite scores in this sample at Time 1), which obscure attempts to detect improvements related to intervention. In this study, total raw scores on the VABS were used as a predictor in regression analyses (excluding motor skills). In group analyses, composite standard scores were presented for comparison with previous research (mean of communication, socialization and daily living skills). However, ratio scores were used for most analyses. Ratio scores for each domain were calculated by dividing age equivalent scores by chronological age and multiplying by 100 (similar to deriving a ratio IQ). Composite ratio scores were very highly correlated with composite standard scores (e.g., r (140) = .93, $p < .001$ at Time 1). However, they provided a wider range of possible outcomes and were less correlated with age at both time points (r (140) = −.58, $p < .001$ versus r (140) = −.79, $p < .001$ at Time 1; r (140) = −.15, p = .071 versus r (140) = −.35, $p < .001$ at Time 2).

2.3.3. Cognitive skills

Information about cognitive functioning was not available at Time 1. At Time 2, cognitive skills were assessed using one of three widely used measures: the Mullen Scales of Early Learning (Mullen, 1995; 74% of children), the Weschler Preschool and Primary Scale of Intelligence (3rd ed.; Weschler, 2002; 25% of children), or the Stanford–Binet Intelligence Scale (4th ed., Thorndike, Hagen, & Sattler, 1986; 1% of children). Different measures were used based on the age and skill level of participants, with younger children and lower functioning older children typically completing the Mullen. Whenever possible, IQ scores were obtained directly from either the Full Scale IQ on the WPPSI or Stanford–Binet or the Early Learning Composite on the Mullen. However, a number of lower functioning children could not attain standard scores. For these children, a ratio IQ was calculated (mental age divided by chronological age, multiplied by 100; as in Perry et al., 2008; Remington et al., 2007).

2.4. Intervention received

2.4.1. IBI group

TPAS provides publicly funded community-based IBI as part of the Ontario IBI Initiative. Formal measures of treatment quality and fidelity were not available. However, service professionals followed program guidelines for the Ontario IBI Initiative (Ministry of Children and Youth Services, 2006). These guidelines stipulate that children must receive behavioral intervention based on the principles of applied behavior analysis, with frequent use of data collection and data-based decision making. In addition, comprehensive curricula must be used, with adaptations to meet individual children's needs. Discrete trial training was combined with more naturalistic teaching approaches, and children were consistently taught to generalize their skills to new contexts. Parents were encouraged to attend training and to meet regularly with treatment staff. For further description of the intervention, see Perry et al. (2008).

Program guidelines mandate that children must receive between 20 and 40 h of IBI per week, except when transitioning to or from treatment. In this sample, prescribed treatment intensity varied from 20 to 35 h/week across children (M = 25.81, SD = 3.44; based on information available for 98% of children). However, children may not have attended all prescribed sessions, and intensity may have been lower at the beginning and end of treatment. Intervention was typically provided in self-contained treatment centers in the community, but was occasionally provided in children's homes.

2.4.2. Waitlist group

Parents of children in the Waitlist group were contacted by telephone to obtain information about the type and quantity of supports received during the waitlist period (available for 94% of children). Based on available data, most children attended school while on the waitlist. Sixty-five percent attended small, specialized classrooms for children with developmental disabilities, 12% attended typical classrooms with support from an educational assistant, and 5% attended typical classrooms without support (18% did not attend school). In addition, many children attended either specialized daycare programs (14%), regular daycare programs with consultation from community agencies (27%), or regular daycare programs with no specialized support (31%; 28% did not attend daycare). On average, children in the Waitlist group attended school and/or daycare for 17.9 h/week (SD = 12.3; range = 0–44 h; only one child attended neither school nor daycare).

In addition to school and daycare services, children on the waitlist accessed interventions such as low intensity behavioral intervention (<10 h/week; 14% of the sample), speech therapy (68% of the sample), occupational therapy (53% of the sample), and behavioral consultation (34% of the sample). Seventy-five percent of parents reported attending training from TPAS and/or another agency (all parents were offered training sessions at TPAS on topics such as addressing challenging behavior and encouraging communication). Lastly, 7% of children took medication related to their autism diagnosis and 14% were on specialized diets or took special supplements during the waitlist period.

3. Results

3.1. Between group comparisons at Time 1 and Time 2

To promote equivalence between the IBI and Waitlist groups at Time 1, a subset of children from the larger dataset described in Section 2.2 were selected for between-group comparisons. Children in the Waitlist group were individually matched to children in the IBI group based on Time 1 age (within 4 months). Matches were possible for all but 2, resulting in 61 matched pairs. Initial age was used to match children because it has been identified as a potentially important predictor of outcome following IBI (e.g., Harris & Handleman, 2000). Although it would have been ideal to match children on other variables in addition to age (e.g., adaptive functioning, autism severity), this would have dramatically reduced sample size, limiting power.

Prior to analyses, all variables were examined for accuracy of data entry and the presence of outliers. VABS scores and IQ estimates were logarithmically transformed, as distributions for these variables were positively skewed. Following transformations, analyses indicated that assumptions regarding normality of the sampling distribution, linearity, homogeneity of variance, homogeneity of regression, and reliability of covariates were adequately met.

Gender did not differ significantly across groups (87% boys in IBI group, 84% boys in Waitlist group, χ^2 (121) = 0.26, $p = .610$), nor did the proportion of children with an autism versus a PDD-NOS diagnosis at intake (54% autism in IBI group, 43% autism in Waitlist group, χ^2 (121) = 1.61, $p = .205$). Mean scores at Time 1 were compared across groups using independent samples *t*-tests. As displayed in Table 1, there were no significant differences in mean scores between groups at Time 1 (all *p* values >.70). However, the interval between test periods (referred to as duration) was longer for the IBI group ($M = 27.84$ months, $SD = 8.11$) than for the Waitlist group ($M = 17.01$, $SD = 2.81$; t (120) = −.9.84, $p < .001$). Due to this difference, children in the IBI group were older for Time 2 assessments (70.77 [6.60] months versus 59.80 [10.58] months, t (120) = −6.87, $p < .001$). As described in the following paragraphs, differences in duration and age at Time 2 were statistically controlled when comparing mean scores at exit.

A series of analyses of covariance were carried out to explore differences between groups at Time 2. Group (IBI or Waitlist) was entered as a fixed factor, and duration and age at Time 2 were entered as covariates. For analyses involving the CARS and the VABS, relevant initial scores were also entered as covariates (cognitive scores at Time 1 were not available to use as a covariate when exploring cognitive outcomes). To better understand the clinical significance of group findings, effect sizes were explored using Cohen's *d* (difference between estimated marginal mean scores at Time 2 divided by pooled standard deviation at Time 2). Effect sizes were considered small if they exceeded 0.2, medium if they exceeded 0.5, and large if they exceeded 0.8 (Cohen, 1992).

Table 2 presents unadjusted means and standard deviations for both groups at Time 2, as well as estimated marginal mean scores. Estimated marginal mean scores are most meaningful to interpret as they correct for duration, age at Time 2, and initial scores (if available). Adjusting for covariates, there was a medium-sized difference in autism severity at Time 2, with milder symptoms observed in the IBI group. Similarly, scores on the VABS were significantly higher for the IBI group across domains (medium effect size). Lastly, there was a large, 19-point, difference in cognitive scores at exit, with results favoring the IBI group. As mentioned in the previous paragraph, it was not possible to control for cognitive scores at Time 1 when exploring cognitive outcomes, as these scores were not available. However, it is unlikely that cognitive skills differed at Time 1, as groups were very similar on all variables at intake, including adaptive skills (adaptive and cognitive skills are highly correlated in children with developmental disabilities: $r = .69$ in Perry et al., 2009; r (140) = .89, $p < .001$ for adaptive composite ratio scores and IQ estimates in this study at Time 2).

Table 1
Group means and standard deviations at Time 1.

	IBI group ($n = 61$)	Waitlist group ($n = 61$)	t	p
Chronological age in months	42.93 (11.53)	42.79 (10.51)	−0.07	.980
Childhood Autism Rating Scale				
Total	32.83 (3.99)	32.62 (3.74)	−0.29	.770
Vineland Adaptive Behavior Scales				
Standard scores				
Composite[a]	55.38 (7.00)	55.49 (7.11)	0.08	.931
Ratio scores				
Composite[a]	30.78 (10.78)	30.79 (10.67)	0.09	.996
Communication	25.47 (15.81)	25.50 (11.97)	0.16	.874
Daily living skills	42.79 (11.97)	42.87 (12.11)	0.03	.977
Socialization	24.08 (9.36)	23.99 (11.22)	−0.12	.900

Note. Scores on the VABS were logarithmically transformed prior to statistical analyses. This table presents untransformed means and standard deviations.
[a] Mean of communication, daily living skills, and socialization scores.

Table 2
Group means, standard deviations, and estimated marginal mean scores at Time 2.

	IBI group ($n = 61$)	Waitlist group ($n = 61$)	F	p	d
Childhood Autism Rating Scale					
Total	30.20 (4.97)	32.57 (5.55)			
	30.00	32.77	4.64	.033	0.53
Vineland Adaptive Behavior Scales					
Standard scores					
Composite[a]	56.34 (14.40)	52.19 (8.77)			
	56.96	50.66	7.20	.008	0.53
Ratio scores					
Composite[a]	41.77 (20.26)	31.15 (11.82)			
	40.75	30.32	9.87	.002	0.63
Communication	46.60 (29.91)	30.33 (16.98)			
	43.45	29.80	7.85	.006	0.56
Daily living skills	44.83 (14.01)	40.03 (11.06)			
	45.04	38.80	5.32	.023	0.49
Socialization	33.90 (19.04)	23.11 (10.85)			
	33.49	21.88	11.77	.001	0.75
Cognitive skills					
IQ estimate	55.80 (26.97)	39.50 (18.93)			
	55.71	36.46	10.32	.002	0.83

Note. This table presents untransformed means and standard deviations for each outcome, followed by estimated marginal mean scores. Estimated marginal mean scores adjust for Time 1 scores (for all but cognitive scores), age at Time 2, and duration, and are transformed back from logarithmic scores when applicable. IQ estimates and scores on the VABS were logarithmically transformed prior to statistical analyses.
[a] Mean of communication, daily living skills, and socialization scores.

At Time 2, 18% of children in the IBI group had IQ estimates in the average range (>85), versus 3.3% in the Waitlist group (χ^2 (121) = 6.97, $p = .008$). Few children had adaptive composite standard scores in the average range at Time 2 (4.9% of IBI group; 0% of Waitlist group, χ^2 (121) = 3.08, $p = .080$). However, 14.8% of the IBI group experienced a large gain of at least 15 points in their adaptive composite score between Time 1 and Time 2 (versus 1.6% in Waitlist group; χ^2 (121) = 6.97, $p = .008$).

3.2. Predictors of outcome

Hierarchical multiple regression was used to predict IQ estimates at Time 2. All 142 children described in Section 2.2 were included in these analyses. Duration was entered at Step 1, as this variable differed for the two groups. At Step 2, the role of initial age was examined controlling for duration. Next, the role of initial adaptive skills was explored, with both age and duration controlled. At Step 4, the role of autism severity was assessed controlling for previously entered variables. Group membership was entered at Step 5 of the model, and interaction variables were entered at Step 6.

Prior to analyses, all predictor variables were centered to reduce multicollinearity. An examination of residual scatter plots indicated that assumptions regarding normality, linearity, and homoscedasticity of residuals were adequately met. Table 3 presents correlations among variables included in these analyses.

Regression results are summarized in Table 4. Although duration and initial age contributed significantly to predictions when they were initially entered into the model, they did not remain significant predictors after information about group membership and interaction variables was added. Controlling for duration and initial age, higher initial adaptive skills contributed a large amount of variance across groups (26%, $p < .001$). When duration, initial age and initial adaptive skill level were controlled, there was a trend ($p = .092$) towards milder initial autism severity contributing an additional 1% to

Table 3
Correlations among variables ($N = 142$).

	1	2	3	4	5	6	7	8	9
(1) Initial age	–	–	–	–	–	–	–	–	–
(2) Initial adaptive skills[a]	.31**	–	–	–	–	–	–	–	–
(3) Initial autism severity[b]	.16*	−.48**	–	–	–	–	–	–	–
(4) Duration	−.56**	−.23**	−.03	–	–	–	–	–	–
(5) Group	.04	.01	.04	.57**	–	–	–	–	–
(6) Group × age	.83**	.29**	.08	−.68**	.00	–	–	–	–
(7) Group × adaptive skills	.31**	.78**	−.36**	−.31**	−.03	.37**	–	–	–
(8) Group × autism severity	.09	−.37**	.76**	−.04	.01	.10	−.48**	–	–
(9) Group × duration	−.67**	−.28**	−.03	.93**	.30**	−.82**	−.37**	−.06	–
(10) Exit IQ estimate	−.30**	.38**	−.41**	.32**	.28**	−.33**	.29**	−.33**	.28**

[a] VABS total raw score excluding motor skills at Time 1.
[b] CARS total score at Time 1.
* $p < .05$.
** $p < .01$.

Table 4
Summary of hierarchical regression analyses for variables predicting IQ estimates at Time 2 ($N = 142$).

Model	Variable	B	$SE\ B$	β	t	p	R^2	ΔR^2	$F\ \Delta$	p
1	Duration	.936	.238	.316	3.938	<.001	.100	.100	15.511	<.001
2	Duration	.632	.283	.213	2.229	.027	.123	.024	3.740	.055
	Initial age	−.388	.200	−.185	−1.934	.055				
3	Duration	.758	.239	.256	3.173	.002	.385	.262	58.744	<.001
	Initial age	−.687	.173	−.328	−3.973	<.001				
	Initial adaptive skills[a]	.641	.084	.539	7.664	<.001				
4	Duration	.774	.237	.261	3.262	.001	.398	.013	2.872	.092
	Initial age	−.583	.183	−.278	−3.193	.002				
	Initial adaptive skills	.545	.101	.458	5.410	<.001				
	Initial autism severity[b]	−.941	.555	−.138	−1.695	.092				
5	Duration	−.023	.327	−.008	−.071	.944	.445	.047	11.639	.001
	Initial age	−.917	.201	−.437	−4.556	<.001				
	Initial adaptive skills	.535	.097	.450	5.512	<.001				
	Initial autism severity	−.940	.535	−.138	−1.757	.081				
	Group	16.201	4.749	.312	3.412	.001				
6	Duration	−1.094	.881	−.369	−1.242	.217	.498	.053	3.494	.010
	Initial age	−.305	.261	−.146	−1.169	.245				
	Initial adaptive skills	.340	.146	.286	2.332	.021				
	Initial autism severity	−1.679	.819	−.247	−2.050	.042				
	Group	25.160	6.222	.485	4.044	<.001				
	Group × duration	.379	.991	.106	.382	.703				
	Group × initial age	−1.327	.408	−.525	−3.255	.001				
	Group × initial adaptive skills	.324	.191	.212	1.697	.092				
	Group × initial autism severity	1.033	1.057	.115	.978	.330				

[a] VABS total raw score excluding motor skills at Time 1.
[b] CARS total score at Time 1.

predictions. Controlling for previously entered variables, group membership accounted for an additional 5% of variance ($p = .001$; group membership would likely have accounted for more variance if it were not so highly correlated with duration, entered previously; between-group comparisons provide more meaningful information about the impact of group membership). Lastly, combined group interaction variables accounted for 5% of unique variance ($p = .010$). As a whole, predictor variables accounted for approximately half of the variability in cognitive outcomes ($R^2 = .50$).

As combined interaction variables contributed significantly to the model, interactions between predictors and group membership were explored in more depth. As displayed in Step 6 of the model, the impact of duration, initial adaptive skill level, and initial autism severity did not vary significantly across groups. However, there was a significant interaction between initial age and group membership. To further investigate this interaction, additional regression analyses were carried out within the two groups separately. Controlling for duration, younger initial age accounted for a moderate amount of variance in outcome for children who had received IBI, R^2 *change* = .13, *F change* (1, 76) = 12.34, $p = .001$. In contrast, initial age did not account for a significant amount of variance in outcome for the Waitlist group, R^2 *change* = .03, *F change* (1, 60) = 1.55, $p = .218$.

4. Discussion

Although several studies have examined the impact of IBI under carefully regulated conditions, few have examined its effectiveness in routine clinical contexts involving children with varied skill levels and family backgrounds. In this study, children receiving IBI were compared to children on a waitlist receiving a range of educational and other support services. Results support the effectiveness of community-based IBI programs. Although groups did not differ significantly at intake, the IBI group had milder autism severity, higher adaptive functioning, and higher cognitive skills at exit. The effect size of group differences was medium for autism severity and adaptive functioning, and large for cognitive skills.

This study provides meaningful information about the types of outcome that could be expected for large, community-based IBI programs treating varied samples of children. Results are generally comparable to those of previous studies exploring the effectiveness of IBI in typical clinical contexts (e.g., Perry et al., 2008; Remington et al., 2007). However, this study reported differences in autism severity and overall adaptive functioning that have not always been found in other community-based research (e.g., Remington et al., 2007).

Although this study supports the effectiveness of IBI, group differences were smaller than those observed in some efficacy studies, with respect to both IQ estimate (e.g., Eikeseth et al., 2007; Howard et al., 2005; Lovaas, 1987) and adaptive functioning (e.g., Cohen et al., 2006; Eikeseth et al., 2007; Howard et al., 2005). Effectiveness studies often report smaller treatment impacts than efficacy studies (Weisz & Jensen, 2001). Smaller treatment benefits may be strongly influenced by factors such as differences in inclusion criteria. In many efficacy studies in the IBI field, children have been under 48 months of age at intake and had ratio IQ scores greater than 35. In contrast, 28% of the children in this study were 4 or 5 years old at intake, and 67% had adaptive composite ratio scores less than 35 at Time 1. Our research team has examined outcomes for a younger, higher functioning sub-sample of children from the dataset in this study, selected using inclusion criteria similar to those employed in many efficacy studies (initial age <48 months, initial adaptive composite ratio score >35; Flanagan, Perry, & Freeman, 2009). In this subsample, differences between groups were larger (e.g., a very large, 32-point difference in IQ estimates at Time 2).

This is the first study to explore predictors of treatment response in IBI relative to a comparison group using multiple regression techniques. Results suggest that initial age is an important predictor of better outcomes in IBI relative to a comparison group, supporting the idea that there may be a period of development during which IBI is most effective at altering developmental trajectories (see Dawson, 2008). A number of previous studies have also reported an association between initial age and outcome in IBI (e.g., Harris & Handleman, 2000; Perry et al., 2011). Studies that have not found this association (e.g., Eikeseth et al., 2007; Lovaas, 1987) have typically involved specific selection criteria, a relatively narrow age range at intake, and a relatively small sample size.

Initial adaptive skill level also emerged as an important predictor in this study, accounting for a large amount of variance when initial age and duration were controlled. However, results suggest that initial adaptive level may be an equally important predictor for children who have received IBI and children who have received other interventions. This is consistent with previous research demonstrating that initial skill level predicts later performance in samples receiving IBI (e.g., Eikeseth et al., 2007; Sallows & Graupner, 2005) and other treatments (e.g., Eaves & Ho, 2004; Gabriels, Hill, Pierce, Rogers, & Wehner, 2001).

Other researchers have reported links between milder initial autism severity and outcome following IBI (e.g., Sallows & Graupner, 2005). This study suggests that autism severity may not play a meaningful role as a predictor when variables such as age and skill level are controlled. Similarly, duration was not a significant predictor of outcome. As suggested by the results of previous research (e.g., Sallows & Graupner, 2005), it is likely that duration plays a different role as a predictor in different subgroups of children receiving IBI, especially when standard scores are used. Some children may continue to make gains on standardized measures with additional years in IBI, up to a certain point, while others may learn new skills but fall further behind same-aged peers.

Combined predictor variables accounted for only half of the variance in cognitive outcomes. It is likely that other variables, such as treatment quality, family involvement, and degree of challenging child behavior also influence the likelihood that children will succeed in IBI. Early response to intervention may be an especially important predictor of good outcomes that was not included in this study. A prospective waitlist controlled study involving the Ontario IBI Initiative that includes additional variables is currently underway (Dunn Geier, Freeman, Perry, & Barrowman, in preparation).

This paper contributes to the literature on outcomes and predictors in the IBI field, by virtue of its size and the fact that it reflects a community-based effectiveness sample of a culturally and socioeconomically diverse group of children with heterogeneous developmental and diagnostic characteristics. However, there are a number of limitations to the study that are important to address. First, as this was a retrospective file review study, clinicians carrying out assessments knew whether children had received IBI. Although data were collected for clinical reasons and clinicians followed best practice guidelines, this knowledge could conceivably have affected results. Secondly, duration between test periods differed significantly across groups. Due to this difference, duration and age at Time 2 had to be controlled statistically when exploring results, and there were strong correlations between duration and group membership in regression analyses. A further limitation of this study is that information about cognitive functioning was not available at Time 1 for either group. Although it is unlikely that cognitive skills differed at intake as groups were virtually equivalent on other measures (see the final paragraph of Section 3.1), this could not be verified. In addition, it would have been ideal to be able to explore the impact of IBI on a broader range of outcomes (e.g., disruptive behavior, parent stress; see Matson, 2007).

A final limitation of this study is that children were not allocated to groups based on random assignment. Some authors have described the considerable difficulties associated with randomization in IBI effectiveness research (e.g., Lord et al., 2005; Remington et al., 2007). Unlike in drug trials, participants cannot be blind to the intervention, and there are ethical challenges associated with changing the standard of care or delaying the provision of service when an intervention is considered to be important to children's success. When randomization is not undertaken, it is ideal to match groups as closely as possible. In this study, parents in both groups selected IBI as a desired intervention for their children, children in both groups were eligible to receive treatment, and children were matched closely based on initial age. Transitions into treatment from the waitlist were made using a standardized waitlist management system that did not include any prioritization other than time of referral. As this was a file review study, volunteer bias could not distort the results.

5. Conclusions

In conclusion, this study suggests that IBI has meaningful effects on a range of outcomes when provided in the context of a large-scale public program serving a heterogeneous group of children. These findings also highlight the importance of ensuring that children begin IBI early, as IBI may be most effective relative to other interventions when children are very young. Findings also suggest that children with higher initial adaptive skills are more likely to experience good outcomes in IBI, although these children may also benefit the most from other interventions.

Older children and children with very low skill levels merit high quality intervention and are likely to benefit from behavioral teaching approaches. However, there may be a skill level above which, or an age below which, children are most likely to make significant gains in IBI relative to other treatments. Additional research is needed to continue to improve our understanding of the most appropriate supports for children with autism at all ages and skill levels.

Acknowledgements

This research was supported by funding from the Regional Autism Programs of Ontario Network. In addition, Helen Flanagan received funding from the Canadian Institutes of Health Research and the Autism Spectrum Interdisciplinary Research Training Program (a CIHR STIHR Program, PI: Holden). We would like to thank the following professionals for assisting with data collection: Sally Bond, Bharathy Kavallappa, Alissa Levy, Kristen McFee, Alice Prichard, Emily Simkins-Strong, Abbie Solish and April Sullivan. In addition, we are grateful for the support of families and staff at the Toronto Partnership for Autism Services (Surrey Place Centre,

Aisling Discoveries, Adventure Place, the Etobicoke Children's Centre, and Geneva Centre for Autism).

References

Bibby, P., Eikeseth, S., Martin, N. T., Mudford, O. C., & Reeves, D. (2002). Progress and outcomes for children with autism receiving parent-managed interventions. *Research in Developmental Disabilities, 23*, 81–104.

Ben-Itzchak, E., & Zachor, D. A. (2007). The effects of intellectual functioning and autism severity on outcome of early behavioral intervention for children with autism. *Research in Developmental Disabilities, 28*, 287–303.

Carter, A. S., Volkmar, F. R., Sparrow, S. S., Wang, J. J., Lord, C., Dawson, G., et al. (1998). The Vineland Adaptive Behavior Scales: Supplementary norms for individuals with autism. *Journal of Autism and Developmental Disorders, 28*, 287–302.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155–159.

Cohen, H., Amerine-Dickens, M., & Smith, T. (2006). Early intensive behavioral treatment: Replication of the UCLA model in a community setting. *Journal of Developmental and Behavioral Pediatrics, 27*, S145–S155.

Dawson, G. (2008). Early behavioral intervention, brain plasticity, and the prevention of autism spectrum disorders. *Development and Psychopathology, 20*, 775–803.

Dunn Geier, J., Freeman, N. L., Perry, A., & Barrowman, N. Multi-site controlled study of the effectiveness of IBI for children with autism in a community based setting, in preparation.

Eaves, L. C., & Ho, H. H. (2004). The very early identification of autism: Outcome to age 4 1/2–5. *Journal of Autism and Developmental Disorders, 34*, 367–378.

Eaves, R. C., & Milner, B. (1993). The criterion-related validity of the Childhood Autism Rating Scale and the Autism Behavior Checklist. *Journal of Abnormal Child Psychology, 21*, 481–491.

Eikeseth, S., Smith, T., Jahr, E., & Eldevik, S. (2007). Outcome for children with autism who began intensive behavioral treatment between ages 4 and 7: A comparison controlled study. *Behavior Modification, 31*, 264–278.

Eldevik, S., Hastings, R. P., Hughes, C., Jahr, E., Eikeseth, S., & Cross, S. (2009). Meta-analysis of early intensive behavioral intervention for children with autism. *Journal of Clinical Child and Adolescent Psychology, 38*, 439–450.

Ferster, C. B., & DeMyer, M. K. (1961). The development of performances in autistic children in an automatically controlled environment. *Journal of Chronic Disorders, 13*, 312–345.

Filipek, P. A., Accardo, P. J., Baranek, G. T., Cook, E. H., Dawson, G., Gordon, B., et al. (1999). The screening and diagnosis of autism spectrum disorders. *Journal of Autism and Developmental Disorders, 29*, 439–484.

Flanagan, H. E., Perry, A., & Freeman, N. L. (2009, May). Comparing apples to apples: Examining the impact of community-based intensive behavioural intervention for children with autism using inclusion criteria employed by model programs. *Poster presented at the 35th annual meeting of the association for behavior analysis international.*

Freeman, N. L., & Perry, A. (2010). Outcomes of intensive behavioural intervention in the Toronto Preschool Autism Service. *Journal on Developmental Disabilities, 16*, 17–32.

Gabriels, R. L., Hill, D. E., Pierce, R. A., Rogers, S. J., & Wehner, B. (2001). Predictors of treatment outcome in young children with autism. *Autism, 5*, 1362–3613.

Goin-Kochel, R. P., Myers, B. J., Hendricks, D. R., Carr, S. E., & Wiley, S. B. (2007). Early responsiveness to intensive behavioural intervention predicts outcomes among preschool children with autism. *International Journal of Disability, Development and Education, 54*, 151–175.

Harris, S. L., & Handleman, J. S. (2000). Age and IQ at intake as predictors of placement for young children with autism: A four- to six-year follow-up. *Journal of Autism and Developmental Disorders, 30*, 137–142.

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., Magiati, I., & Charman, T. (2009). Systematic review of early intensive behavioral interventions for children with autism. *American Journal on Intellectual and Developmental Disabilities, 114*, 23–41.

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*, 695–708.

Lovaas, O. I. (1967). Behavior therapy approach to the treatment of childhood schizophrenics. In *Minnesota symposia on child psychology* (vol. 1, pp. 108–159). Minneapolis, MN: University of Minneapolis Press.

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., Koegel, R., Simmons, J. P., & Long, J. S. (1973). Some generalization and follow-up measures on autistic children in behavior therapy. *Journal of Applied Behavior Analysis, 6*, 131–166.

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*, 308–812.

Makrygianni, M. K., & Reed, P. (2010). A meta-analytic review of the effectiveness of behavioural early intervention programs for children with autistic spectrum disorders. *Research in Autism Spectrum Disorders, 4*, 577–593.

Matson, J. L. (2007). Determining treatment outcome in early intervention programs for autism spectrum disorders: A critical analysis of measurement issues in learning based interventions. *Research in Developmental Disabilities, 28*, 207–218.

Ministry of Child and Youth Services. (2006, August). *Autism intervention program guidelines* Retrieved from http://www.abacuslist.ca/Client/ASO/ABA.nsf/object/MinistryAutismGuidelines/$file/AutismGuidelines-en.pdf.

Mullen, E. M. (1995). *Mullen scales of early learning*. Circle Pines, MN: American Guidance Service.

National Autism Center. (2009). *National standards report*. Randolph, MA: National Autism Center.

New York State Department of Health. (1999). *Autism/pervasive developmental disorders: Clinical practice guidelines technical report*. New York: Author.

Perry, A. (2002). Intensive early intervention program for children with autism: Background and design of the Ontario preschool autism initiative. *Journal on Developmental Disabilities, 9*(2), 121–128.

Perry, A., Condillac, R. A., Freeman, N. L., Dunn Geier, J., & Belair, J. (2005). Multi-site study of the Childhood Autism Rating Scale (CARS) in five clinical groups of young children. *Journal of Autism and Developmental Disorders, 35*, 625–634.

Perry, A., Cummings, A., Dunn Geier, J., Freeman, N. L., Hughes, S., LaRose, L., et al. (2008). Effectiveness of intensive behavioral intervention in a large, community-based program. *Research in Autism Spectrum Disorders, 2*, 621–642.

Perry, A., Cummings, A., Dunn Geier, J., Freeman, N. L., Hughes, S., Managhan, T., et al. (2011). Predictors of outcome in children receiving intensive behavioral intervention in a large, community-based program. *Research in Autism Spectrum Disorders, 5*, 592–603.

Perry, A., Flanagan, H. E., Dunn Geier, J., & Freeman, N. L. (2009). Brief report: The Vineland Adaptive Behavior Scales in young children with autism spectrum disorders at different cognitive levels. *Journal of Autism and Developmental Disorders, 39*, 1066–1078.

Reichow, B., & Wolery, M. (2009). Comprehensive synthesis of early intensive behavioural interventions for young children with autism based on the UCLA Young Autism Project Model. *Journal of Autism and Developmental Disorders, 39*, 23–41.

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.

Sallows, G. O., & Graupner, T. D. (2005). Intensive behavioural treatment for children with autism: Four-year outcome and predictors. *American Journal on Mental Retardation, 110*, 417–438.

Schopler, E., Reichler, R. J., & Renner, B. R. (1988). *The Childhood Autism Rating Scale (CARS)*. Los Angeles, CA: Western Psychological Services.

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.

Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1984). *Vineland Adaptive Behavior Scales*. Circle Pines, MN: American Guidance Service.

Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005). *Vineland Adaptive Behavior Scales* (2nd ed.). Circle Pines, MN: American Guidance Service.

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Stanford–Binet Intelligence Scale* (4th ed.). Itasca, IL: Riverside Publishing.

Wechsler, D. (2002). *Wechsler preschool and primary scale of intelligence-III*. San Antonio, TX: Psychological Corporation.

Weisz, J. R., & Jensen, A. L. (2001). Child and adolescent psychotherapy in research and practice contexts: Review of the evidence and suggestions for improving the field. *European Child and Adolescent Psychiatry, 10*, I12–I18.

Yoder, P., & Compton, D. (2004). Identifying predictors of treatment response. *Mental Retardation and Developmental Disabilities, 10*, 162–168.