

Table of Contents

SECTION B: A Quantitative Examination of the Sustainability of the Better Beginnings, Better Futures Project Based on School Measures

A.	TEACHER RATINGS IN THE YOUNGER CHILD SITES	B-1
	SAMPLE	B-1
	The Research Sample Who Attended Junior Kindergarten in 1998-1999	B-2
	The Research Sample Who Attended Senior Kindergarten (SK) in 2002-2003	B-3
	MEASURES	B-3
	Junior Kindergarten Teacher Ratings	B-3
	Senior Kindergarten Teacher Ratings	B-4
	Scale Overlap Between the JK and SK Teacher Ratings	B-5
	ANALYSES	B-6
	RESULTS	B-7
	Analyses of JK and SK Teacher Ratings on Children in the Younger Child Sites in 1999 and 2003	B-7
	Hyperactivity-inattention	B-8
	Emotional-anxiety disorder	B-9
	Physical aggression	B-10
	Prosocial behaviour	B-11
	Social competence	B-12
	Summary of analyses of JK and SK teacher ratings on children in the younger child sites in 1999 and 2003	B-13
	Analyses of SK Teacher Ratings on Children in 2003	B-14
	Summary of analyses of SK teacher ratings on children in 2003	B-18
B.	TEACHER RATINGS IN THE OLDER CHILD SITES	B-18
	SAMPLE	B-18
	The Research Sample Who Attended Grade 3 in 1997-98	B-18
	The Research Sample Who Attended Grade 3 in 2002-2003	B-19
	MEASURES	B-20
	Emotional and Behavioural Problems	B-20
	Social Skills	B-20
	Grade Repetition	B-20
	Special Education Service Use	B-20

ANALYSES	B-21
RESULTS	B-21
Results of Analyses of Grade 3 Children’s Emotional and Behavioural Problems	B-22
Hyperactivity-inattention	B-22
Emotional-anxiety disorder	B-23
Physical aggression	B-24
Indirect aggression	B-25
Delinquency	B-26
Summary of analyses of Grade 3 children’s emotional and behavioural problems	B-27
Results of Analyses of Grade 3 Children’s Social Skills	B-28
Conflict management	B-28
Cooperation	B-29
Assertiveness	B-29
Summary of analyses of Grade 3 children’s social skills	B-30
Results of School Performance Measures	B-30
Special education involvement	B-30
Grade repetition	B-31
C. SPECIAL EDUCATION USE: A LONGITUDINAL NEIGHBOURHOOD-LEVEL ANALYSIS	B-32
ANALYSES	B-33
RESULTS	B-33
Younger Child Sites	B-33
Older Child Sites	B-35
CONCLUSIONS	B-36

As described in Section A of this report, the Better Beginnings programs have developed and matured over the past few years. Each of the Better Beginnings project sites have “fine-tuned” their programs, dropped programs that were not meeting the needs of the community, and added new programs. The cohort of children and families who have participated in the research since 1994 reflect experiences with programs that in some cases were undergoing revisions and modifications based on what seemed to the program decision-makers in each community to be working and not working. It is of substantial interest and value to document program outcomes, now that the programs are “proud”, stably funded and project personnel are more confident about what is working in their communities.

For this research, a cross-sectional design was employed, collecting information about children living in the Better Beginnings and comparison neighbourhoods at a fixed age of the child, namely the upper age range of the Better Beginnings programs. In the younger child sites, where Better Beginnings programs focus on children from birth to 4 years of age, information was collected from Senior Kindergarten (SK) teachers in the spring of 2003. Teachers rated school readiness, health and behavioural problems for students in their SK class. This provided an assessment of children who had spent a good deal of their preschool years in Better Beginnings communities, and allowed for an ongoing determination of the sustainability of program effects through comparisons to data collected previously in 1999, on 4-year old children in the younger child Better Beginnings and comparison communities.

A similar cross-sectional strategy was employed in the older child sites where Better Beginnings programs focus on children from Junior Kindergarten (JK) through Grade 2. Information was collected at Grade 3 in 2003 on child development indicators using teacher ratings of children’s behaviour, special education class placements and grade retention. In addition, information was collected from the Ontario Ministry of Education’s Principals October Report database on special education use in all schools in the Better Beginnings and comparison neighbourhoods. Analyses of these measures were carried out between Better Beginnings and comparison sites and then were contrasted with the same measures that had been collected on Grade 3 children in 1998. This allowed for an assessment of changes in Grade 3 outcome effects over time in the older child Better Beginnings and comparison site neighbourhoods.

A. TEACHER RATINGS IN THE YOUNGER CHILD SITES

SAMPLE

In order to determine whether Better Beginnings, Better Futures is having an effect on children growing up in a “Better Beginnings” neighbourhood, research teams obtained teacher ratings of children living in the five younger child project sites and a comparison neighbourhood on a group of children attending JK in the spring of 1999 and a group of children attending SK in the spring of 2003.

The Research Sample Who Attended Junior Kindergarten in 1998-1999

Children who entered Junior Kindergarten in the fall of 1998 in the five younger child Better Beginnings project sites represented the first cohort of children to move through the full four years of Better Beginnings programming from birth to age 4 in the years 1994 to 1998.

Some of the children attending JK in 1998 were part of the Better Beginnings longitudinal research sample. In the five younger child Better Beginnings project sites and one comparison site, all children who were born in 1994 were eligible to participate in the longitudinal research. The most widely used method of recruitment involved the support of hospitals at which mothers from our study areas were likely to give birth. With consent from the mother, records department staff passed on the names of new mothers to Better Beginnings staff, who could then explain the study and, with the mother's agreement, arrange an interview. If Better Beginnings did not have enough staff to cover all the hospitals at which mothers at a site were likely to give birth, arrangements were made with Public Health Units to send material about the study to those who were eligible from their lists of new mothers, who could then contact Better Beginnings for more information, or send in a consent form directly. Some site research groups also found it useful to visit prenatal classes or to leave the same type of information sent out by Public Health Units with organizations likely to be in contact with mothers of young children, who could let the mothers know about the research and, with consent, pass their names to the Better Beginnings site researchers. Recruitment of children born in 1994 was also done through local Better Beginnings programs, such as playgroups, so that each year from 1994 to 1998, new children and their families were added to the longitudinal research group. We can compare participation levels to the estimated number of eligible families. For the younger child sites, we can employ 1996 Census data to estimate the number of children aged 0-4, then divide by 5 to estimate the number in a single-year cohort. In the first year of data gathering for the longitudinal samples, the participation rate was 58%.

Because three of the five schools located within the Guelph Better Beginnings project site did not offer JK, JK teacher ratings for this site cannot be reported. In the remaining four project sites and comparison site at JK, there were 556 children in the longitudinal research sample. Sixty-six percent of those children (365/556) were still living in either one of the Better Beginnings project sites or comparison site when they attended JK. We received parental consent for the JK teachers to complete a teacher rating form on 71% (258/365) of children still living in a Better Beginnings project or comparison site. One hundred and seventy-eight children lived in one of the four Better Beginnings project neighbourhoods (Kingston, $n = 43$; Ottawa, $n = 16$; Regent Park, $n = 56$; Walpole Island, $n = 63$), and 80 children were living in the Peterborough comparison neighbourhood.

Our researchers also requested parents' permission to gather teacher ratings on the rest of the 1998-1999 JK class (these children were living in a Better Beginnings site or comparison site, but were not part of our longitudinal research sample). Teacher ratings were provided for 202 children in three project sites (Kingston, $n = 86$; Ottawa, $n = 90$; and Regent Park, $n = 26$), and for 232 children from the comparison neighbourhood. We did not receive permission to obtain teacher ratings from the non-research children attending JK on Walpole Island, but our research sample represented approximately 66% of the JK class.

When the JK teacher ratings from the longitudinal research sample and their classmates are combined, the total JK sample size is as follows: 380 children in four project sites (Kingston, $n = 129$; Ottawa, $n = 106$; Regent Park, $n = 82$; Walpole Island, $n = 63$), and 312 children from the comparison neighbourhood.

The Research Sample Who Attended Senior Kindergarten (SK) in 2002-2003

Our research team obtained teacher ratings of SK children living in one of the five younger child sites or living in one of two comparison neighbourhoods (Hamilton and Peterborough) in the spring of 2003. The Hamilton comparison neighbourhood was added to the research design in 2001 to improve comparability with the project sites on employment, income and ethnic composition.

This group of SK children were born in 1997, three years after the longitudinal sample described above. The SK teacher ratings provided a “snapshot” of how children were doing in both the project sites and comparison (control group) sites. We chose a SK sample rather than a JK because of the national trend to assess school readiness at SK using the Early Development Instrument (described in detail below).

Since the identity of the SK children remained anonymous to the researcher, many principals gave teachers permission to fill out ratings for the children in her/his class without involving the parents. Otherwise, principals required a consent form to be sent home to the parents. In total, teacher ratings were provided for 516 children in five project sites (Guelph, $n = 91$; Kingston, $n = 88$; Ottawa, $n = 133$; Regent Park, $n = 170$; Walpole Island, $n = 34$), and for 397 children from the comparison neighbourhoods. This sample of 913 children represents approximately 90% of all eligible SK children from the project and comparison neighbourhoods.

Of the 913 children that SK teachers completed the EDI, 15% (141/913) had not attended Junior Kindergarten. Over 40% of children who did not attend JK were from the Guelph site (61/141) because the public school board did not offer JK. We decided to exclude these 141 children from the analyses for two reasons: (1) the EDI measure is a school readiness measure, so the children who did not attend JK are likely to score lower on this measure; (2) the Guelph site had a disproportionately larger sample of children who did not attend JK, and therefore, the results from this site might be negatively biased. Therefore, for the analyses of the SK EDI measure, the trimmed sample was 772 (Guelph, $n = 30$; Kingston, $n = 79$; Ottawa, $n = 117$; Regent Park, $n = 156$; Walpole Island, $n = 34$; comparison neighbourhoods $n = 356$).

MEASURES

Junior Kindergarten Teacher Ratings

Appraisal for Better Curriculum (ABC) is a Toronto Board of Education (1990) program developed to ensure that the abilities, self-concepts, learning styles and experiences of all children are considered when programs are planned. Better Beginnings, Better Futures used its appraisal form, consisting of 17 items which attempts to assess school readiness (see Appendix F for a list of items). In the spring of 1999, it was completed by Junior Kindergarten (JK) teachers

for each student in their class who was participating in the Better Beginnings research project and living in a Better Beginnings site or comparison site. A total score was created by combining ratings on the 17 items; scores for each item could range from 1 indicating the skills or behaviours were not yet apparent to 5 indicating the skills or behaviours were extremely well developed. The internal consistency of this scale, using Cronbach alpha, was 0.97.

JK teachers were also asked to assess children's prosocial behaviour, using the *prosocial scale* from the Canadian National Longitudinal Survey of Children and Youth (NLSCY; 1995)¹. The *prosocial scale* consists of ten items and answers for each item include 0 (never), 1 (sometimes) and 2 (often), yielding a total score from 0 to 20, with high scores indicating frequent prosocial behaviours. Three types of children's emotional and behavioural problems were also examined based on the NLSCY: *emotional-anxiety disorder* (8 items), *physical aggression* (6 items), and *hyperactivity-inattention* (7 items). Each item was rated as "never or not true" (0), "sometimes true" (1), or "often or very true" (2). A description of each of these scales can be found in Appendix F. The internal consistency of these four scales using Cronbach alpha is as follows: prosocial scale 0.92; emotional-anxiety disorder 0.91; physical aggression 0.93; and hyperactivity-inattention 0.94.

Senior Kindergarten Teacher Ratings

The Early Development Instrument (EDI) was developed at the Canadian Centre for Studies of Children at Risk at McMaster University to assess readiness to learn at school among 5 year-old children, prior to their entry to Grade 1 (Janus & Offord, 2000). The EDI was created in order to provide communities with a feasible, acceptable and psychometrically sound instrument that can be used for whole populations of children to monitor community efforts to improve early years outcomes over time (Janus & Offord, 2000).

Readiness to learn at school was defined as the child's ability to meet the task demands of school and to benefit from the educational activities that are offered by the school. The EDI combines several areas that have been identified as relevant to children's school readiness (Doherty, 1997; Kagan, 1992) including: *physical health and well-being*; *social competence*; *emotional maturity*; *language and cognitive development*; and *communication skills and general knowledge*. The scores on each of the five domains are calculated by taking the average score on a range of responses. Cases with more than 25% missing values for any domain do not receive a valid score for that domain. Domain scores on each of the five scales could range from 0 to 10, with higher scores indicating more of that ability (see Appendix G for a description of the five domains.)

The internal consistency of the five domains was explored using Cronbach alpha (Janus, 2004). All five domains showed very satisfactory internal consistency levels: physical health and well-being 0.84; social competence 0.96; emotional maturity 0.92; language and cognitive development 0.93; and communications skills and general knowledge 0.95. Janus (2004) also

¹ Many scales and questions used in the Better Beginnings research are the same as those used in the NLSCY. This was done in order to allow comparison of measures with a nationally representative sample of Canadian children of the same age. For information on the development and psychometric properties of the NLSCY questions and scales, see Statistics Canada and Human Resource Development Canada (1995).

explored the test-retest reliability of the EDI, inter-rater reliability, and the validity of the measure; interested readers are referred to her paper for more details.

The *physical health and well-being* domain refers to the child's physical preparedness for the school day, fine and gross motor skills, energy level throughout the day, and physical independence; there are 13 questions in total. The *social competence* domain refers to the child's competence and cooperation in working together with others, ability to remember and follow rules, curiosity and eagerness, approaches to learning and problem-solving; there are 26 items in total. The *emotional maturity* domain covers prosocial behaviour, aggression, inattention and hyperactivity, and anxious behaviours; there are 30 items. The *language and cognitive development* domain refers to the child's ability to use the language correctly and covers cognitive development aspects of language and numeracy, in several areas using 26 items: basic literacy and numeracy skills, interest and memory, and more complex literacy. The *communication skills and general knowledge* domain covers the child's ability to clearly communicate one's own needs and thoughts in a way that is understandable to both adults and other children, the ability to understand others, clear articulation, as well as aspects of general knowledge; there are 9 questions in this domain.

Scale Overlap Between the JK and SK Teacher Ratings

We were interested in examining whether there is a difference over time in children's behaviours and readiness to learn between children living in a neighbourhood where Better Beginnings programs were offered compared to children living in a neighbourhood without Better Beginnings programs. We compared teachers' ratings of children in both the project and comparison site in the spring of 1999 (when the research children were in JK) and four years later in the spring of 2003 (when the research sample of children was in SK). As described above, different questionnaires were used at JK vs. SK. However, many of the items from the JK questionnaire and the SK EDI measure were drawn from the same NLSCY survey.

We began with the four NLSCY scales that we used at JK and determined how many of the same items within each scale were also used at SK. Nine of the ten items on the *prosocial scale* were used at both JK and SK. Five of the eight *emotional-anxiety disorder* scale were used at both JK and SK. Three of the six *physical aggression* items were used at both JK and SK. Finally, six of the seven items of the *hyperactivity-inattention* scale were used at both points in time. (See Appendix H for a description of these scales.) The creation of each subscale was based on the examination of attenuated reliability coefficients and expected correlation of two sets of items. The first set contained the items that were used in both data sets, and the second set had the items that were left out from the base data set. For example, emotional-anxiety disorder scale is an NLSCY scale and it is created by utilizing eight items. It is available in JK data (what we call here the base data) and its reliability alpha is 0.91. On the contrary the SK EDI teacher report form has only five items that could be matched with the eight JK items. So we took the base sample, calculated the attenuated alpha coefficients for five items (0.87), and then for three items (alpha=0.80). From this, we calculated the observed and expected correlations. Finally we took the ratio of observed and expected correlations; if the ratio was close to 1, we concluded that the five-item emotional-anxiety subscale would fairly represent the full eight-item scale. Scale scores on each of the four scales could range from 0 to 10, with higher scores indicating more of that behaviour.

We also examined the relationship between the EDI social competence domain asked at SK and the ABC measure used at JK. With SK as our base data set, we identified seven items from the social competence domain that were the same or similar in the JK sample. Following the procedure mentioned above, we created the seven-item subscale of social competence. The correlation between this subscale and ABC measure was very high (0.97), which led us to take one of the two measures. We picked the social competence subscale for subsequent analysis. We called this subset of items the *social competence* subscale which had scores ranging from 0 to 10, with higher scores indicating higher social competence (see Appendix H for a description of the scale).

ANALYSES

The first step of analysis was examination of descriptive statistics for each outcome variable gathered from the JK and SK teachers. The results were expressed in terms of means and standard deviations of sites separately. Each of the project sites was separately contrasted against the comparison sites. Also the combined sample of Better Beginnings project sites was compared with the comparison sites.

Analysis of variance (ANOVA) techniques (also known as *omnibus* test) were employed to test for an overall experimental effect. ANOVA allowed us to perform planned contrasts² to compare between groups of interest without inflating the Type I error rate (the probability of accidentally finding a difference to be statistically significant when it isn't really so). Each of the younger child project sites was compared with the Peterborough comparison site, such as Guelph vs. Peterborough and Kingston vs. Peterborough. One way to compare them would have been to carry out four t-tests, which would inflate the *familywise error rate* (where each and every t-test is an opportunity to make a Type I error). This can be avoided by carrying out planned contrasts in the context of ANOVA. This approach breaks down the variance accounted for by the model into component parts and then tests the null hypothesis that there is no difference in means between two components of interest (e.g. Guelph vs. Peterborough). Thus, the omnibus F-ratio provides the effect of overall experimental manipulation and the contrast F-ratios provide information about which groups were affected.

For each outcome variable, a 3-way Univariate ANOVA was carried out, where the dependent variable was the outcome variable in question and the independent variables were Better Beginnings site, time and sex of child. We wanted to see whether or not the Better Beginnings site means improved from time 1 to time 2 when compared with comparison site means, holding the sex of child as a control variable. We included sex of child as a control variable for two reasons. First, numerous studies suggest that sex of child works as a strong differentiating predictor of the child's emotional and behavioural problems when measured as we have measured them (Offord & Lipman, 1996; Willms, 2002). Second, this is the only trustworthy socio-demographic that was consistently available from teachers for these data. The interaction between time and site is the crucial element of our analysis that shows whether the overall

² There are several kinds of contrasts (e.g. linear, quadratic, orthogonal, non-orthogonal, etc.), and we used linear orthogonal contrast. See Rosenthal and Rosnow (1985).

change in means from time 1 to time 2 differs from site to site. We carried out linear orthogonal contrast in order to detect the specific group differences of interest to us.

Because the sites were not chosen randomly, statistical inference cannot be made beyond them. Use of a post-randomization test allowed an estimate of whether random processes working only within the samples gathered could plausibly have created differences between sites of the kind that were observed. The tests were conducted with 15,000 random replications. Although these tests typically provided smaller standard errors than standard ANOVA output, the results were consistent, so, in the interests of wider intelligibility, we have reported standard ANOVA significance test results.

When many variables are analyzed, measured on differing scales or in different units, readers are often aided by the conversion of the results into a common measure of effect size. Whereas p values express the probability of a particular outcome being due to chance, the effect size provides an indication of how large the outcome is, taking into consideration the variability in the measure in question. Following Cohen's (1977) recommendations, effect sizes were calculated for all variables on which significant interactions have emerged. Under the two-point in time design, the effect size is just the time 2 score minus time 1 score, under the accepted model, divided by the standard deviation for the time 1 sample. For dichotomous variables, where we are interested in changing percentages, they are transformed, under the accepted model, by the formula

$$\Phi = 2 \arcsin p^5$$

where p is the percentage of interest. The difference between Φ for time 2 and Φ for time 1 is then taken.

Although Cohen (1977) pointed out that to call an effect large or small was somewhat arbitrary, and that the use of terms of this kind should vary from topic to topic, the conventions he suggested are widely used in psychological research, and were employed here. Cohen's threshold effect sizes, and their labels, are as follows: 0.20 small; 0.50 moderate; and 0.80 large. In other fields, these conventions do not apply as well, because a change that is not common may be of great importance. Because there are no cross-disciplinary standards for what effects should be considered important, effect sizes were calculated for all significant interaction effects.

RESULTS

Analyses of JK and SK Teacher Ratings on Children in the Younger Child Sites in 1999 and 2003

The analyses reported in this section examined differences in teacher ratings between groups of JK children in the Better Beginnings and comparison sites collected at the end of the project demonstration phase in 1999, and differences in similar ratings collected in the same sites four years later on a group of SK children in 2003. Teacher ratings on five measures of children's behaviour were analyzed: hyperactivity-inattention, emotional-anxiety disorder, physical aggression, prosocial behaviour, and social competence. The average (mean) ratings for children

from each of the Better Beginnings and comparison sites, along with standard deviations, samples sizes and results of 2-way analyses of variance (ANOVA) are presented in Appendix I for each of the five measures. If the difference between the Better Beginnings and comparison sites was more favourable for the Better Beginnings sites in the 2003 ratings than they had been in 1999, this was considered to be a positive outcome for the sustainability of Better Beginnings programs. Alternatively, if the difference in teacher ratings between Better Beginnings and comparison sites was less favourable for the Better Beginnings sites in 2003 than in 1999, this was considered a negative outcome for the sustainability of Better Beginnings programs. These differences are reflected in the ANOVA results as the interaction between the “site” variable (Better Beginnings site vs. comparison site) and the “time” variable (1999 vs. 2003).

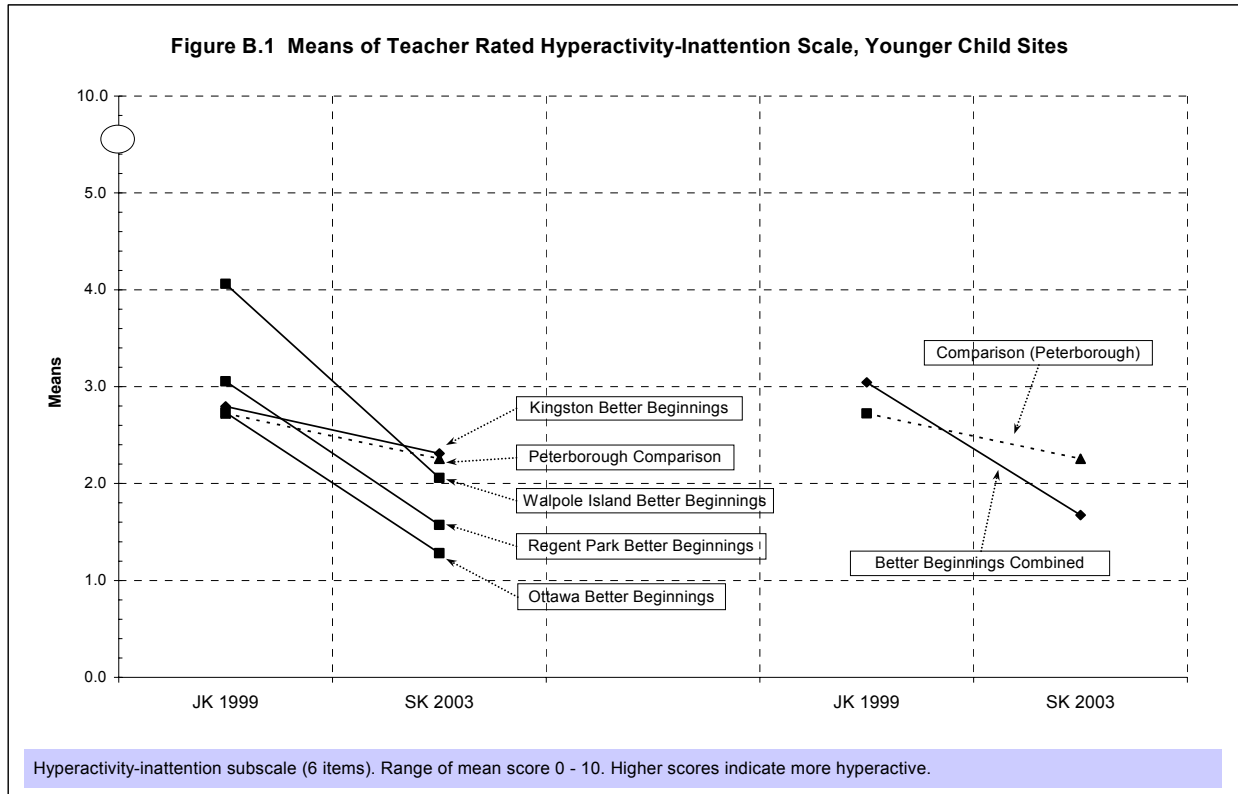
The results of the site by time interaction effects presented in Appendix I are summarized in Table B.1. This table indicates whether the interaction effects are positive or negative for each of the four Better Beginnings sites in Kingston, Ottawa, Regent Park and Walpole Island vs. the comparison site in Peterborough, and for the four Better Beginnings site scores combined vs. the Peterborough comparison site. Also included in Table B.1 is the probability that each outcome is statistically significant, and, if significant, the effect size for the interaction. (Since most four year old children in the Guelph Better Beginnings site did not attend JK in 1999, no JK teacher ratings for this site were available for these analyses.)

Table B.1 1999 vs. 2003 Teacher Ratings in the Younger Child Sites: Summary of Analyses

JK/SK Scales	Kingston vs. Peterborough	Ottawa vs. Peterborough	Regent Park vs. Peterborough	Walpole Island vs. Peterborough	4 Project Sites vs. Peterborough
Hyperactivity-Inattention	+	+* E.S.=.32	+* E.S.=.31	+* E.S.=.55	+** E.S.=.29
Emotional-Anxiety Disorder	-.** E.S.=.45	- ^M	-	+	-
Physical Aggression	-	+** E.S.=.36	-	-* E.S.=.41	+
Prosocial	-	+** E.S.=.38	+	+*** E.S.=.77	+
Social Competence	-.** E.S.=.45	+*** E.S.=.58	+	-	+

* p<.05, ** p<.01, *** p<.001, M p<.10, E.S. = Effect Size

Hyperactivity-inattention. The results in Table B.1 indicate that the teacher ratings of children’s hyperactivity-inattention behaviours showed a significantly positive interaction effect for the Better Beginnings sites combined vs. the comparison site (p<.01; E.S.=.29). Ratings of these behaviours were also significantly positive for the individual Better Beginnings sites in Ottawa (p<.05; E.S.=.32), Regent Park (p<.05; E.S.=.31) and Walpole Island (p<.05; E.S.=.55) vs. the comparison site in Peterborough. The average teacher ratings of children’s hyperactivity-inattention are presented in Figure B.1.

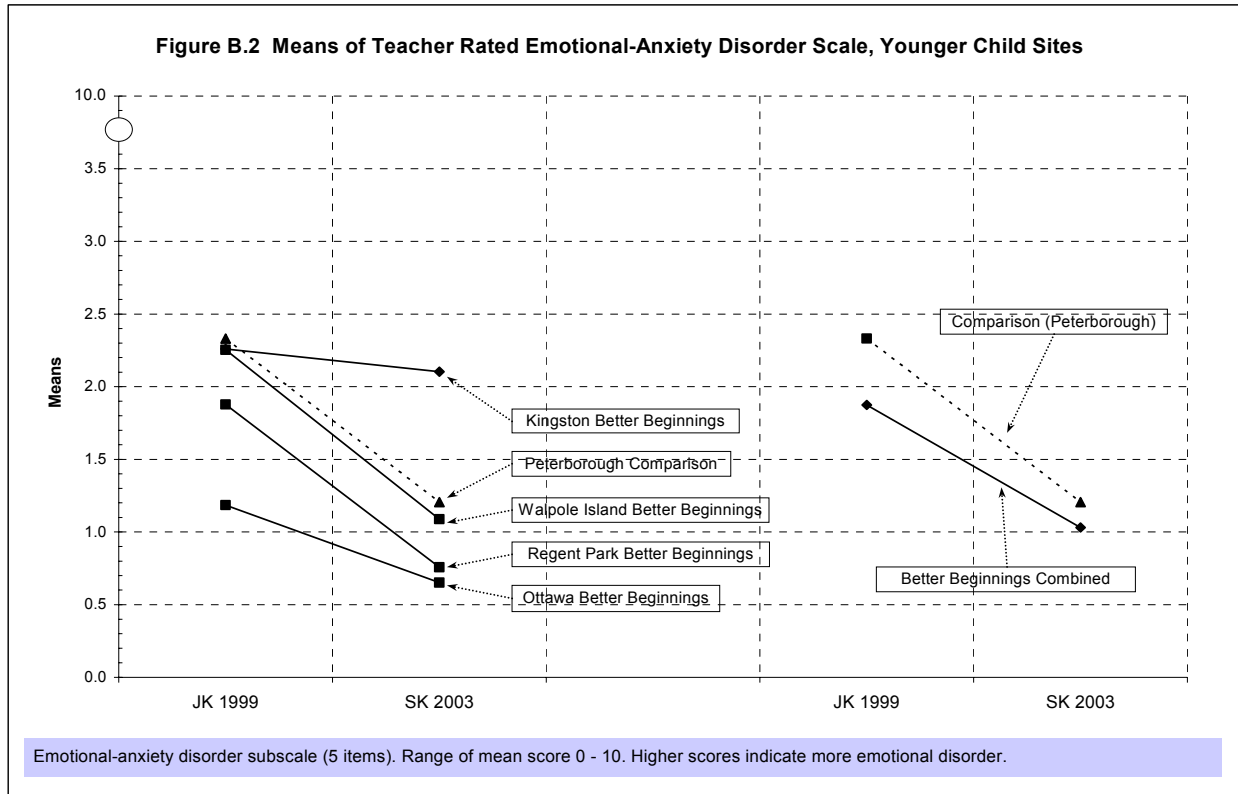


The first thing apparent in Figure B.1 is that the average teacher rating scores were lower for the SK children in 2003 than they were for the JK children in 1999, reflecting a developmental decrease in these behaviours in children from four to five years of age.

The significant positive effect for the combined Better Beginnings sites vs. the comparison site is portrayed on the right side of Figure B.1, and resulted from the fact that the average decrease in hyperactivity scores from 1999 to 2003 for children from the Better Beginnings sites was greater than the decrease over that period for children from the comparison site.

The average ratings from the individual sites, presented on the left side of Figure B.1, reveal larger decreases in teacher ratings of hyperactivity-inattention behaviours from 1999 to 2003 in Ottawa, Regent Park and Walpole Island Better Beginnings sites relative to the Peterborough comparison site, while the decreases over this period in the Kingston Better Beginnings site exactly parallel those in Peterborough.

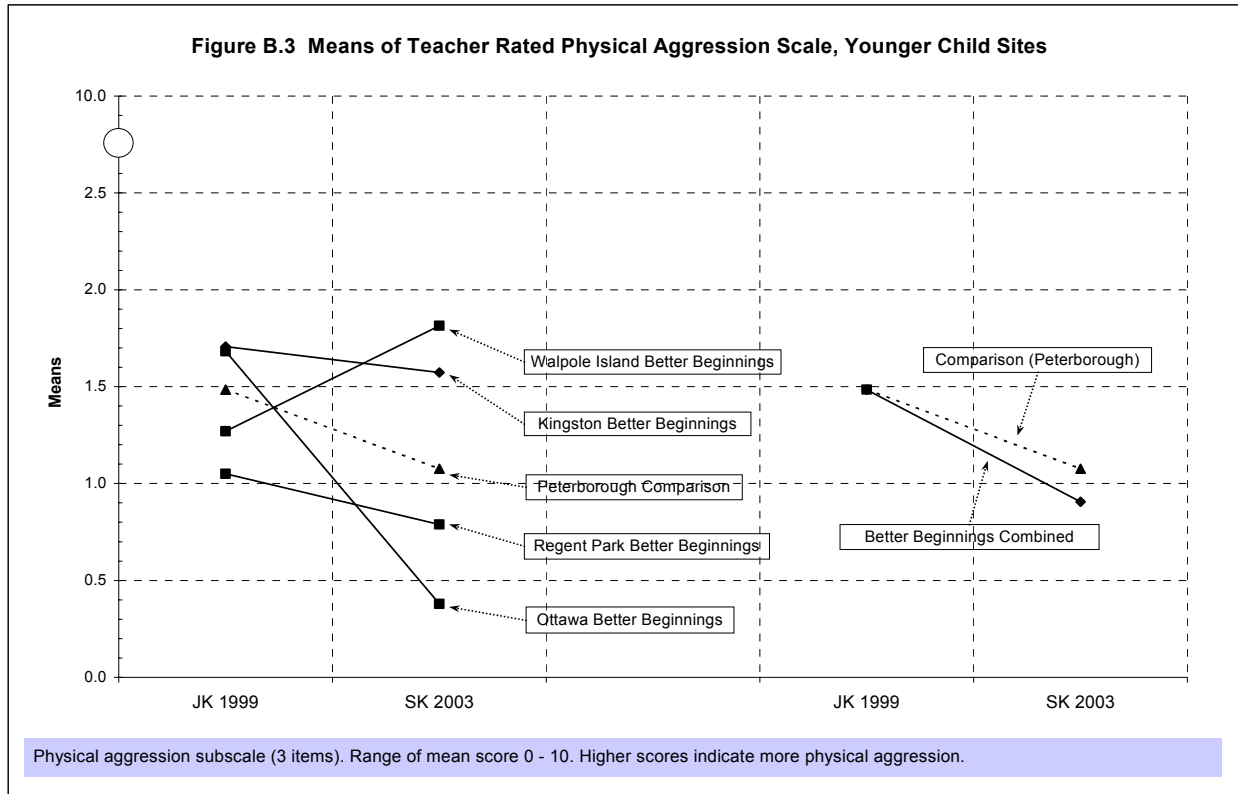
Emotional-anxiety disorder. Average teacher ratings on the emotional-anxiety disorder scale, reflecting children's anxious and depressed behaviours are presented in Figure B.2.



As was the case in ratings of hyperactivity-inattention behaviours, there was a general decrease from the 1999 ratings of JK children to the 2003 ratings of SK children, likely reflecting a general developmental decline in these behaviours in children from 4 to 5 years of age. The only significant interaction between site and time was a negative effect for the Kingston Better Beginnings site vs. the Peterborough comparison site ($p < .01$; $E.S. = .45$). In the 1999 JK ratings, both sites showed the same level of teacher-rated emotional disorder. However, at SK in 2003, ratings in Kingston were substantially higher than in Peterborough, yielding the significant negative interaction effect for Kingston presented in Table B.1.

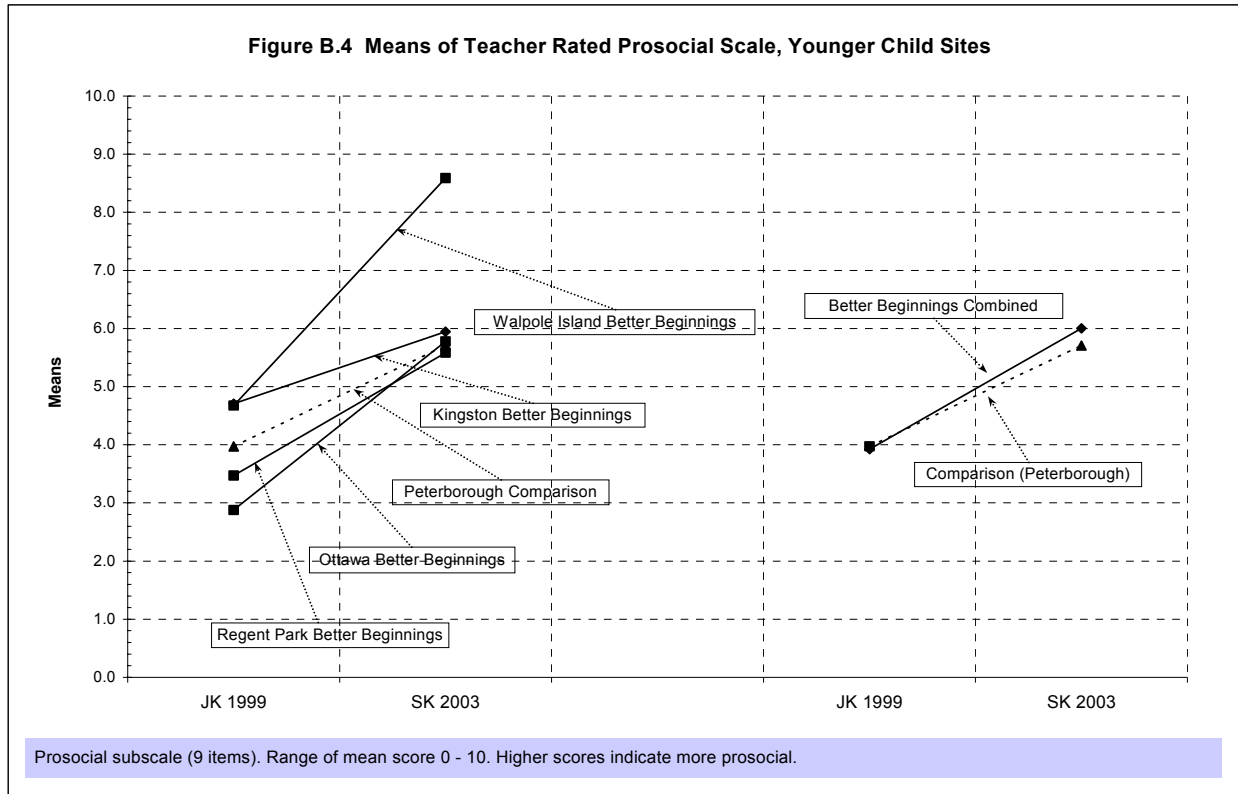
As shown on the right side of Figure B.2, there was no significant change in the average difference between the combined Better Beginnings sites and the comparison site in 1999 vs. 2003.

Physical aggression. The average teacher ratings of children’s physical aggression are presented in Figure B.3, with all sites but Walpole Island showing lower ratings in 2003 in SK ratings than in the 1999 JK ratings.



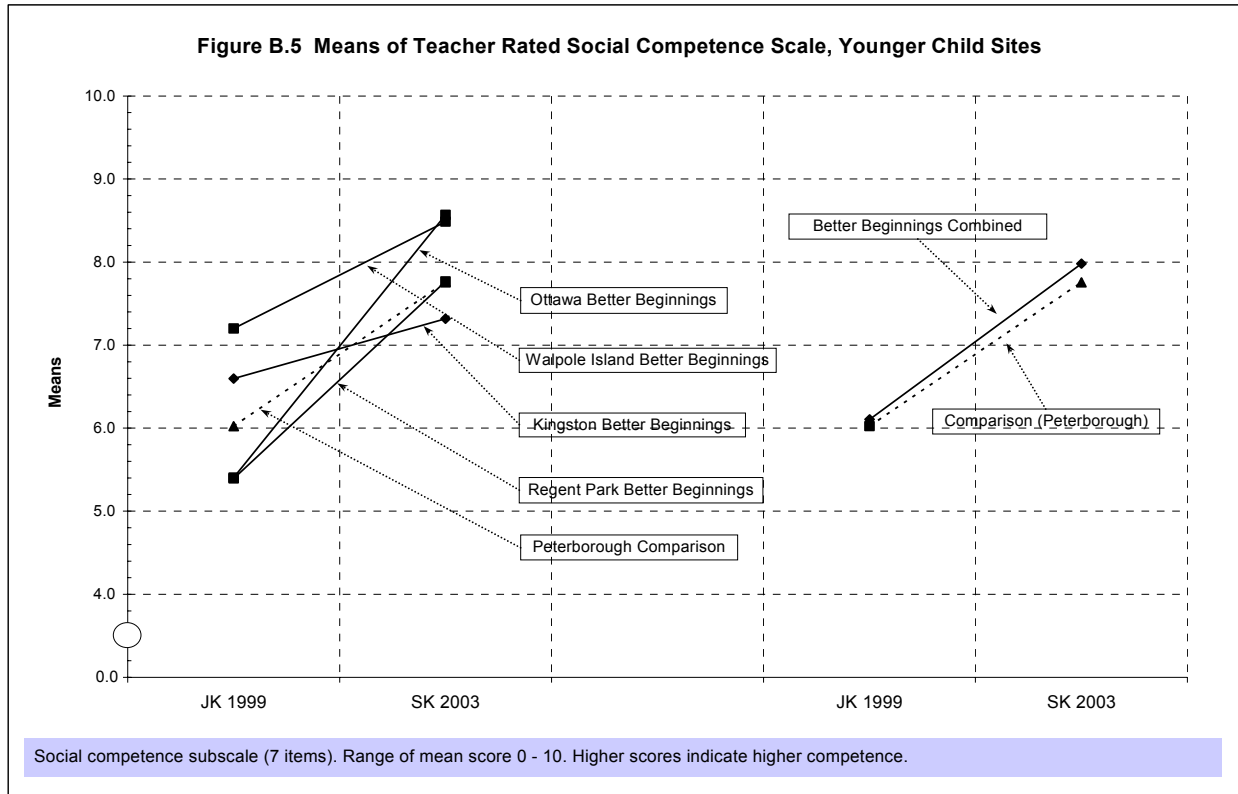
As presented in Table B.1, analyses revealed two significant interactions between a Better Beginnings site vs. the Peterborough comparison site in 1999 vs. 2003. Teacher ratings from the Ottawa Better Beginnings site showed a much larger decrease from 1999 to 2003 than the Peterborough comparison site ($p < .01$; $E.S. = .36$), while the reverse pattern appeared in the teacher ratings from Walpole Island which increased from 1999 to 2003 while Peterborough showed a small decrease ($p < .05$; $E.S. = .41$).

Prosocial behaviour. The overall trend of teacher ratings, as presented in Figure B.4, is a substantial increase in prosocial behaviours from JK to SK students, presumably reflecting the better social interaction skills in the older children.



In terms of site differences, as presented in Table B.1, students from both the Ottawa ($p < .01$, $E.S. = .38$) and Walpole Island ($p < .001$, $E.S. = .77$) site showed significantly greater increases in average teacher ratings from 1999 to 2003 than those for the Peterborough comparison site.

Social competence. As with the prosocial measure just described, the average teacher ratings of children’s social competence presented in Figure B.5 showed consistently higher scores in SK than in JK, again presumably reflecting the greater social maturity of 5 year olds.



Two significant interaction effects between Better Beginnings sites and the Peterborough site from JK to SK resulted (see Table B.1). Although ratings by teachers in the Kingston Better Beginnings site were higher at JK in 1999 than in Peterborough, this difference was reversed in the 2003 SK ratings, resulting from a smaller increase from JK to SK in Kingston than the increase in Peterborough ($p < .01$; $E.S. = .45$).

The opposite result occurred when comparing Ottawa with Peterborough. Ottawa JK students were rated by their teachers as showing lower average prosocial behaviour than JK students in Peterborough in 1999, but by 2003, the SK ratings by Ottawa teachers were substantially higher than those in Peterborough, resulting in the significant positive interaction effect ($p < .001$; $E.S. = .58$).

Summary of analyses of JK and SK teacher ratings on children in the younger child sites in 1999 and 2003. A major goal of the Better Beginnings, Better Futures project is to reduce emotional and behavioural problems in young children and foster more prosocial behaviour and social skills.

In this section, ratings of hyperactivity-inattention, emotional-anxiety problems, physical aggression, prosocial behaviour and social competence that had been collected from JK teachers in four of the five younger child Better Beginnings sites (Kingston, Ottawa, Regent Park, Walpole Island) and one non-project comparison site (Peterborough) in 1999 at the end of the project demonstration phase, were compared to ratings on these same five measures collected from SK teachers in the same schools in 2003. Analyses focused on whether differences

between children living in the Better Beginnings neighbourhoods and those living in the non-project comparison neighbourhood had increased, decreased or remained stable over the four-year period.

For teacher ratings of hyperactivity-inattention, three of the four Better Beginnings sites showed larger decreases from 1999 to 2003 than the Peterborough comparison site, while the fourth Better Beginnings site in Kingston paralleled the change in Peterborough. When the ratings of the four Better Beginnings sites were combined, their combined average score in 2003 was substantially lower than that in Peterborough, suggesting that the Better Beginnings neighbourhood programs strengthened between 1999 and 2003 in reducing children's hyperactive behaviour in school.

Several patterns of differences among the Better Beginnings sites are apparent from the summary of analyses presented in Table B.1. Ottawa showed the most positive pattern of results, with significantly greater improvements relative to the Peterborough comparison site in four of the five measures, the exception being the measure of emotional-anxiety disorder. Kingston, on the other hand, showed significantly smaller improvements over the four-year period relative to Peterborough on two of the five measures: emotional-anxiety disorder and social competence. When combined, the scores from the four Better Beginnings sites showed larger improvements from 1999 to 2003 than those from Peterborough on four of the five measures, although statistically significant only for the measure of hyperactivity-inattention.

Analyses of SK Teacher Ratings on Children in 2003

The average (mean) teacher ratings of SK children on the five EDI domains are presented in Table B.2 for each Better Beginnings site, the five Better Beginnings sites combined and the comparison sites. Also, for reference purposes, the average teacher ratings on each domain from an Ontario normative sample of 24,682 SK children in 2003 are presented.

Table B.2 Average (Mean) EDI SK Teacher Ratings (2003)

EDI Domains	Guelph	Kingston	Ottawa	Regent Park	Walpole Island	5 Better Beginnings Sites Combined	Comparison Sites	Ontario Provincial Average
Physical Health and Well-Being	8.1 -	8.4 -	8.9*** + E.S.=.39	8.4 -	9.0* + E.S.=.45	8.6 +	8.5	8.5
Social Competence	7.4 ^M -	7.7 -	8.6* + E.S.=.27	8.1 +	8.6 +	8.2 +	8.0	8.5
Emotional Maturity	7.0** - E.S.=.48	7.4* - E.S.=.24	8.4** + E.S.=.34	8.2* + E.S.=.19	8.5* + E.S.=.39	8.0 +	7.8	8.1
Language & Cognitive Development	8.4 +	7.9 -	8.0 +	8.7*** + E.S.=.33	8.2 +	8.3 ^M +	8.0	8.5
Communication Skills and General Knowledge	7.6 +	7.5 +	7.8* + E.S.=.25	7.2 -	9.1*** + E.S.=.87	7.6 ^M +	7.3	7.8

* p<.05, ** p<.01, *** p<.001, M p<.10, E.S. = Effect Size

In Table B.2, if a Better Beginnings site average score is higher than the comparison sites' score for a particular domain, the Better Beginnings average is assigned a (+); if it is lower than the comparison sites' average, it is assigned a (-). If a difference reaches statistical significance, the p value and effect size is presented. The average ratings from each of the Better Beginnings sites and comparison sites, along with standard deviations, sample size and results of ANOVA are presented in Appendix J.

As shown in Table B.2, teacher ratings on the first EDI domain, Physical Health and Well Being, were significantly higher in both the Ottawa (p<.001; E.S.=.39) and Walpole Island Better Beginnings sites (p<.05; E.S.=.45) relative to the comparison sites' average. For the EDI domain of Social Competence, SK teachers from the Ottawa Better Beginnings site rated their students, on average, significantly higher than did those from the comparison sites (p<.05; E.S.=.27). Also, the average scores for both the Ottawa and Walpole Island Better Beginnings sites were slightly higher than the Provincial average on both the Physical Health and Well-Being Scale and the Social Competence Scale.

For ratings of Emotional Maturity, both Guelph (p<.01; E.S.=.48) and Kingston average ratings (p<.05; E.S.=.24) were significantly lower than the comparison sites, while the three Better Beginnings sites in Ottawa (p<.01; E.S.=.34), Regent Park (p<.05; E.S.=.19), and Walpole Island (p<.05; E.S.=.39) all showed average teacher ratings that were significantly higher than the comparison sites, as well as being higher than the Ontario Provincial average.

The next two domains of the EDI, Language and Cognitive Development and Communication Skills and General Knowledge reflect the skills most frequently associated with early primary school success (La Paro & Pianta, 2000; Rimm-Kaufman, 2004). As shown in Table B.2, four of

the five Better Beginnings sites showed higher average ratings on both of these EDI domains than the comparison sites. For Language and Cognitive Development only the Kingston average was slightly below that for the comparison sites. Regent Park’s average was significantly higher than the comparison sites ($p < .001$; $E.S. = .33$), and when the scores from the five Better Beginnings sites were combined, the resulting average was higher than the comparison sites, a difference that was marginally significant ($p < .10$). For the Communication and General Knowledge EDI scale, only the Regent Park Better Beginnings site showed an average score just slightly below the comparison sites, while the other four Better Beginnings sites’ averages were above the comparison sites. Both the Ottawa ($p < .05$; $E.S. = .25$) and Walpole Island ($p < .001$; $E.S. = .87$) scores were significantly higher, and the higher average score for the five Better Beginnings sites combined was marginally significant compared to the comparison sites ($p < .10$).

In summary, the average EDI SK teacher ratings of SK students in 2003 were generally higher in the Better Beginning sites than the comparison sites on all 5 EDI domain scales and marginally significant on two. Results for EDI ratings in the Ottawa and Walpole Island Better Beginnings sites were consistently and significantly higher than the comparison site schools on most of the five domains, while Regent Park showed significantly higher scores than the comparison sites on two of the five scales for Emotional Maturity and Language/Cognitive Development. The only significantly lower scores occurred for the Emotional Maturity measure in both the Guelph and Kingston Better Beginnings sites.

In order to explore site differences further on the teachers’ ratings of children’s school readiness, the percentage of children whose ratings placed them below the 10th percentile of the EDI scores for the Province of Ontario was calculated separately for each site and for each of the five EDI domain measures. These percentages are presented in Table B.3.

Table B.3 Percentage of Children with EDI Domain Scores below the Ontario Provincial 10th Percentile

EDI Domains	Guelph	Kingston	Ottawa	Regent Park	Walpole Island	5 Better Beginnings Sites Combined	Comparison Sites
Physical Health and Well-Being	36.7 -	27.8 -	9.4** +	16.7 +	5.9* +	17.3 +	20.8
Social Competence	26.7 -	25.3 -	11.2 +	8.3* +	9.1 +	13.8 +	16.6
Emotional Maturity	26.7 -	26.6 -	6.9** +	9.0* +	12.1 +	13.5 +	17.4
Language & Cognitive Development	13.3 +	19.2 -	12.9 +	5.1** +	14.7 +	11.4 +	14.9
Communication Skills and General Knowledge	16.7 +	14.3 +	12.0* +	12.2* +	2.9** +	12.1** +	20.5

Peterborough and Hamilton Comparison Sites data have been combined.

* $p < .05$

** $p < .01$

A score below the Provincial 10th percentile on an EDI domain is seen as indicative of a child having serious problems in that aspect of school readiness.

Differences between the percentages for each Better Beginnings site and the comparison sites were analyzed, and the results of these analyses are summarized in Table B.3 by indicating whether the percentage for each Better Beginnings site as well as for the five Better Beginnings sites combined was higher (-) or lower (+) than for the comparison sites (for a full description of the Chi-Square values and sample sizes, please see Appendix K).

In the Guelph Better Beginnings site, there was a higher percentage of children showing serious readiness problems than in the comparison sites for 3 of the 5 EDI domains and in the Kingston site for 4 of the 5 EDI domains. None of these differences, however, was statistically significant. For the Ottawa, Regent Park, and Walpole Island Better Beginnings sites, as well as for the five Better Beginnings sites combined, the percentage of serious problems was lower than the comparison sites for all five EDI domains. These differences in the direction of fewer serious problems in the Better Beginnings sites were statistically significant for 3 of the EDI domains in Ottawa, 4 in Regent Park, and 2 in Walpole Island. The most consistent result occurred in the measure of Communication and General Knowledge, with all five Better Beginnings sites showing a lower percentage of children with serious problems than occurred in the comparison sites.

It should also be noted that the percentages of children with serious problems in the Ottawa, Regent Park and Walpole Island Better Beginnings sites are generally near or below the 10% average for the Province as a whole, while these percentages for the Guelph and Kingston Better Beginnings sites, as well as for the comparison sites, are consistently higher than 10%.

Another method of EDI analysis that is used for identifying the degree of serious difficulties being experienced by SK children in a particular school or geographic area is the percentage of children who fall below the normative provincial 10th percentile score in “one or more” of the five EDI domains. These percentages for the Better Beginnings and comparison sites, as well as for the five Better Beginnings sites combined and for the 2003 Ontario normative sample are presented in Table B.4. (See Appendix L for full details of these analyses.)

Table B.4 Percentage of Children Falling Below the Ontario 10th Percentile Score in One or More of the EDI Domains

Guelph	Kingston	Ottawa	Regent Park	Walpole Island	All Better Beginnings Sites Combined	Comparison Sites	Ontario
53%	47%	26% *	29%	26%	33%	38%	24%
-	-	+	+	+	+		

* p<.05 using Fisher’s Exact Test (2 sided)

The picture that emerges from Table B.4 is that there is a very high percentage of children rated by their teachers as experiencing serious difficulties in one or more of the EDI domains in the Guelph and Kingston Better Beginnings sites relative to the comparison sites as well as the Ontario average of 24%. Approximately half of the children rated by SK teachers in each of these two Better Beginnings sites showed extremely low EDI scores in one or more area of school readiness. The comparable percentages from the Ottawa, Regent Park and Walpole Island Better Beginnings sites are lower than the comparison sites and close to the Ontario average.

Summary of analyses of SK teacher ratings on children in 2003. The teacher ratings of SK children on the EDI in 2003 present a picture of generally better functioning in the Better Beginnings sites than the comparison sites, supporting the view that the Better Beginnings programs in the younger child sites continue to have a positive impact on children's functioning five years after the end of the projects' demonstration phase in 1998. The only negative effects occurred in the Guelph and Kingston Better Beginnings sites, where teacher ratings of children's Emotional Maturity were significantly lower than the comparison sites, and these two sites also showed poorer performance in the areas of Social Competence and Physical Health. This appears to have resulted from a generally higher percentage of children in both Guelph and Kingston who were rated by their teachers as having serious difficulties in these three areas of school readiness as reflected in EDI scores below the 10th percentile for the Province.

The EDI results were most consistently positive in the Ottawa, Regent Park and Walpole Island Better Beginnings sites and also on scores that combined all five of the Better Beginnings sites. Teacher ratings on the two EDI domains of a) Learning and Cognitive Development, and b) Communication Skills and General Knowledge showed the most positive results in favour of the five Better Beginnings sites relative to the comparison sites.

B. TEACHER RATINGS IN THE OLDER CHILD SITES

SAMPLE

In order to determine whether differences continue to exist in early primary school children growing up in a "Better Beginnings" neighbourhood, our research team obtained teacher ratings of children living in the three older child project sites (Cornwall, Highfield and Sudbury) and two comparison sites (Etobicoke and Ottawa-Vanier) in the spring of 1998 and 2003.

The Research Sample Who Attended Grade 3 in 1997-98

Children who entered Grade 3 in the fall of 1997 in the three older child Better Beginnings project sites represented the first cohort of children to move through the full four years of Better Beginnings programming from ages 4 to 8 in the years 1993 to 1997.

Some of the children attending Grade 3 in 1997-1998 were part of the Better Beginnings longitudinal research sample. This sample of children and their families were recruited in 1993 when the children were enrolled in Junior Kindergarten. A total of 554 children were involved in the research from JK to Grade 3. Two hundred and fifty-five were in one of the three Better

Beginnings project neighbourhoods (Cornwall, $n=66$; Highfield, $n=77$; Sudbury, $n=112$), and 299 were in one of two comparison neighbourhoods (Etobicoke, $n=117$; and Ottawa-Vanier, $n=182$) which were selected based on similar demographic characteristics to the three Better Beginnings project sites.

The research sample in the three Better Beginnings project sites represented 50 – 60% of the entire birth cohort of children in their neighbourhood, based on school records. Sample bias was checked by asking teachers to rate all children in their classrooms, then comparing ratings for children in the research sample and those outside it. These analyses were carried out for each year of data collection from 1993/4 to 1997/8 and no significant differences in teacher ratings of mental health problems or social skills were found. Sample attrition for the five year period was 7.8%, and analyses of measures comparing children and families who dropped out of the research sample from those who completed all four years of data collection yielded no consistent differences between the two groups.

As of Grade 3, 81% of those children (451/554) were still living in either one of the Better Beginnings project sites or a comparison site. We received parental consent for the Grade 3 teachers to complete a teacher rating form on 87% (394/451) of children still living in a Better Beginnings project or comparison site. One hundred and sixty-four children lived in one of the three Better Beginnings project sites and 230 in one of the two comparison neighbourhoods.

Our researchers also requested parents' permission to gather teacher ratings on the rest of the 1997-1998 Grade 3 class (these children were living in a Better Beginnings site or comparison site, but were not part of our longitudinal research sample). Teacher ratings were provided for 113 children in two of the three project sites (Highfield, $n=61$; and Sudbury, $n=52$), and for 270 children from the comparison neighbourhoods. Principals' permission to gather teacher ratings on the non-research children in the Cornwall schools was not obtained.

When the Grade 3 teacher ratings from the longitudinal research sample and their classmates are combined, the total Grade 3 sample size is as follows: 277 children in three project sites (Cornwall, $n=53$; Highfield, $n=110$; Sudbury, $n=114$), and for 500 children from the comparison neighbourhoods (Etobicoke, $n=240$; and Ottawa-Vanier, $n=260$).

The Research Sample Who Attended Grade 3 in 2002-2003

Our research team obtained teacher ratings of Grade 3 children living in one of the three older child sites or comparison neighbourhoods (Etobicoke and Ottawa-Vanier) in the spring of 2003. This group of children were born five years later than the longitudinal sample described above, so the Better Beginnings programs may have developed and matured since 1997-98. These teacher ratings provided a “snapshot” of how children were doing in both the project sites and comparison (control group) sites.

Since the identity of the Grade 3 children remained anonymous to the researcher, many principals gave teachers permission to fill out ratings for the children in her/his class without involving the parents. Otherwise, principals required a consent form to be sent home to the parents. In total, teacher ratings were provided for 312 children in the three project sites (Cornwall, $n=99$; Highfield, $n=128$; Sudbury, $n=85$), and for 541 children from the comparison

neighbourhoods (Etobicoke, $n = 285$; Ottawa-Vanier, $n=256$). This sample represents approximately 92% of all eligible Grade 3 children from the project and comparison neighbourhoods.

MEASURES

Emotional and Behavioural Problems

Teachers rated children on five types of children's emotional and behavioural problems: hyperactivity-inattention, emotional-anxiety disorder, physical aggression, indirect aggression and delinquency. These five types of childhood problems were selected because they cover a range of emotional and behavioural factors, and because they were measured with the same scales used to collect data from teachers in the NLSCY (1995), therefore allowing comparison with a nationally representative sample (Offord & Lipman, 1996; Willms, 2002). Each item was rated as never (0), sometimes (1) or often (2). Appendix M describes items for each of these five scales. The *hyperactivity-inattention* scale consists of seven items with a Cronbach alpha reliability of .93. The *emotional-anxiety disorder* scale consists of eight items with a Cronbach alpha reliability of .89. The *physical aggression* scale consists of six items with a Cronbach alpha reliability of .93. The *indirect aggression* scale consists of five items with a Cronbach alpha reliability of .95. Finally, the *delinquency* scale consists of three items with a Cronbach alpha reliability of .79.

Social Skills

Teachers were asked to rate the children the Social Skills Rating Scale (SSRS) (Gresham & Elliott, 1990). The SSRS consists of descriptions of a wide variety of prosocial skills and behaviours (e.g. "the child makes friends easily", "the child controls his or her temper when arguing with other children", and "the child is self-confident in social situations such as parties or group outings"). Each behaviour is rated on a 3-point scale with a score of 0 (never), 1 (sometimes), or 2 (very often). The SSRS yielded three subscale scores: conflict management, cooperation, and assertiveness. The *conflict management* subscale consists of six items and has a Cronbach alpha reliability of .89. The *cooperation* subscale consists of nine items and has a Cronbach alpha reliability of .93. The *assertiveness* subscale consists of eight items and has a Cronbach alpha reliability of .84. Appendix M describes items for each of these three subscales.

Grade Repetition

Teachers were asked whether the child had ever repeated a grade. If the teacher answered "no", the child received a score of "0". If the teacher answered "yes", the child received a score of "1".

Special Education Service Use

The assessment of special education service use was based on questions from the teacher report form: a) the child having an individual education plan, b) being identified by an Identification, Placement and Review Committee, and c) receiving any services related to his or her exceptionality (e.g., resource teacher, therapist). The answers to these three questions were

combined to create a binary variable called *Received any Special Education/Service*. If the teacher answered “no” to all three questions, the child received a score of “0”. If the teacher answered “yes” to one or more of the three questions, the child received a score of “1”.

ANALYSES

The analyses reported in this section examined differences in teacher ratings between a group of Grade 3 children in the three older child Better Beginnings project sites in Cornwall, Highfield and Sudbury, and their respective comparison sites in Ottawa-Vanier and Etobicoke collected in 1998 at the end of the project demonstration period and differences in the same teacher ratings between a group of Grade 3 children in the same sites five years later in 2003. Teacher ratings on three aspects of children’s functioning collected and analyzed: 1) emotional and behavioural problems; 2) social skills and prosocial behaviour; and 3) school performance.

The data analyses procedures were the same as employed in the JK (1999) – SK (2003) comparisons reported earlier for the younger child sites. Using three-way ANOVA procedures, differences between Better Beginnings and comparison sites in 1998 were compared with differences between the same sites in 2003. If the differences between the Better Beginnings and comparison sites was more favourable for the Better Beginnings sites in the 2003 ratings than they were in 1998, or equally favourable in both sets of ratings, this was considered a positive (+) outcome for the sustainability of the Better Beginnings programs. Alternatively, if the difference in teacher ratings between the Better Beginnings and comparison sites was less favourable for the Better Beginnings sites in 2003 than in 1998, this was considered a negative (-) outcome for the sustainability of the Better Beginnings programs over the five year period. These differences are reflected in the ANOVA results as the interaction between the “site” variable (Better Beginnings vs. comparison sites) and the “time” variable (1998 vs. 2003).

ANOVA is not an appropriate choice for analyzing a dichotomous outcome variable. We have two outcome variables from the Grade 3 teacher data that are dichotomous for which we have used hierarchical loglinear technique in the analysis. A hierarchical loglinear analysis is a statistical technique for the analysis of a frequency data in multiway cross-tabulation tables such as a 2X2X2 contingency table. It represents a three-way table with three variables, each with two levels (e.g. learning disability – yes/no, site – project/comparison, and time – time1/time2). The 3-way interaction was analyzed by fitting the model and observing the overall goodness-of-fit as assessed by the likelihood ratio chi-square statistic (Knoke & Burke, 1980).

RESULTS

The results of the ANOVA are presented for each measure in Appendix N along with means, standard deviations, percentages, and sample sizes for each analysis. These results are summarized in Table B.5 where a (+) sign indicates a more favourable difference for the Better Beginnings site than the comparison site in 2003 relative to the children five years earlier in 1998, while a (-) means a Better Beginnings effect weakened, became smaller or was reversed from 1998 to 2003. Also included in Table B.5 is the probability that each outcome is statistically significant, and if significant, the effect size for the interaction.

Table B.5 Better Beginnings, Better Futures Older Child Sites: Summary of 1998 vs. 2003 Grade 3 Teacher Data Analyses

Teacher Ratings	Cornwall vs. Ottawa-Vanier Comparison Site	Sudbury vs. Ottawa-Vanier Comparison Site	Highfield vs. Etobicoke Comparison Site	Better Beginnings Sites Combined vs. Comparison Sites Combined
Emotional & Behavioural Problems:				
Hyperactivity-Inattention	+* E.S.=.42	+* E.S.=.27	-	+* E.S.=.21
Emotional-Anxiety Disorder	+	+	-	+
Physical Aggression	+ ^M	+	+	+* E.S.=.26
Indirect Aggression	+* E.S.=.46	+	-	+
Delinquency	+* E.S.=.37	+* E.S.=.35	+	+** E.S.=.25
Social Skills:				
Conflict management	+* E.S.=.46	+** E.S.=.49	+	+*** E.S.=.41
Cooperation	+	+	+	+* E.S.=.23
Assertiveness	-	+ ^M	-	+
School Performance:				
Received Special Education/Services	+**	+*	+	+***
Grade Repetition	+	-	+	+

* p<.05, ** p<.01, *** p<.001, M p<.10, E.S. = Effect Size

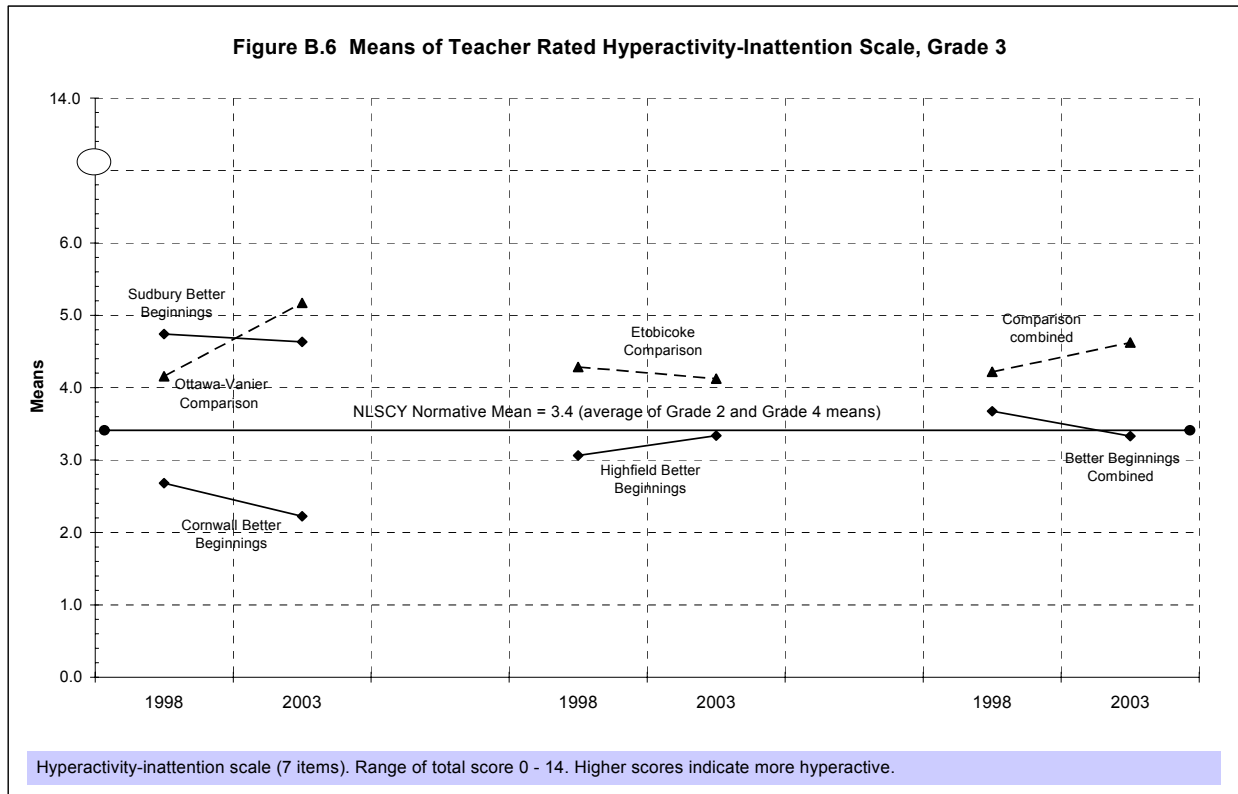
The overall picture of the results summarized in Table B.5 is one of general improvement in the Better Beginnings sites relative to the comparison sites. This is especially true for the Cornwall and Sudbury Better Beginnings sites, and for the analyses of the combined Better Beginnings sites scores relative to the combined comparison site scores. Highfield, on the other hand, showed no significant effects either positive or negative. The results of the analyses for each of the 10 measures summarized in Table B.5 are described in more detail in the following sections.

Results of Analyses of Grade 3 Children’s Emotional and Behavioural Problems

Teacher ratings of the five different types of child emotional and behavioural problems employed in the NLSCY were analyzed: hyperactivity-inattention, emotional-anxiety disorder (anxious and depressive behaviours), physical aggression, indirect aggression, and delinquency.

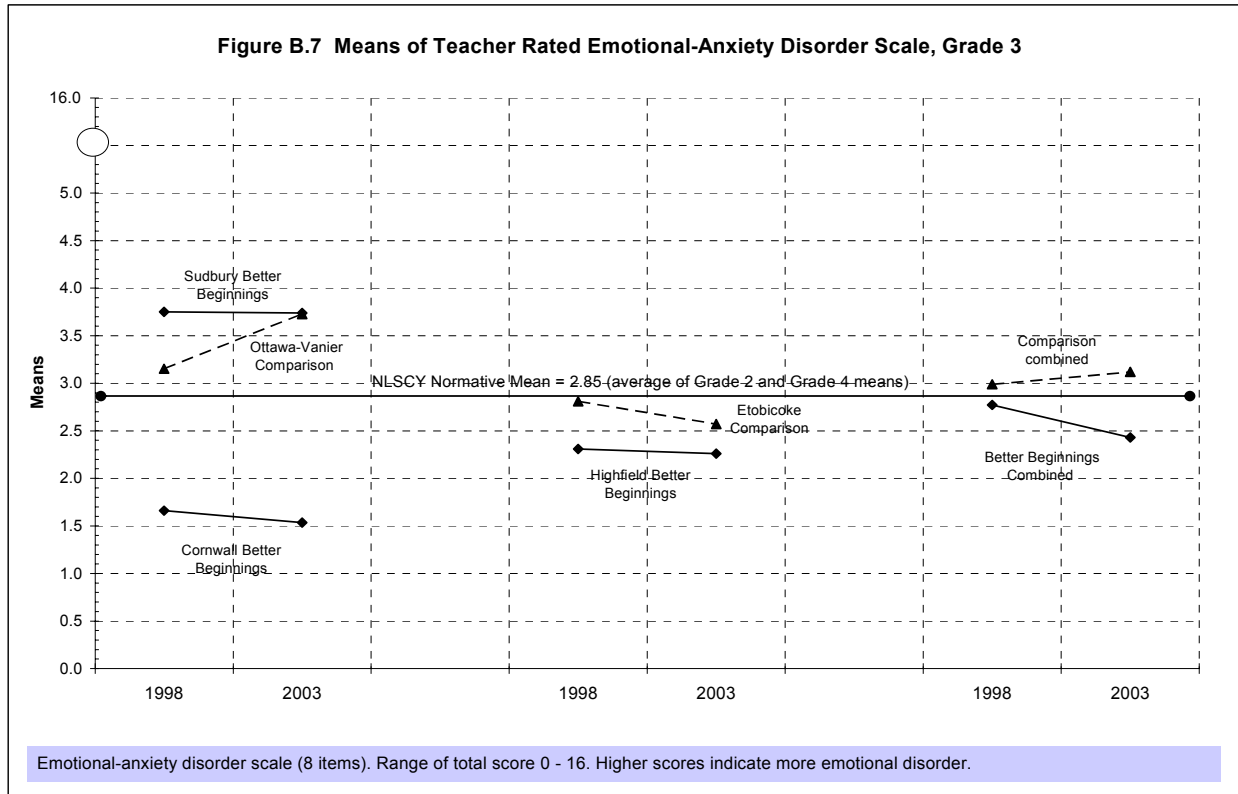
Hyperactivity-inattention. The means for teacher ratings on the hyperactivity-inattention scale are presented in Figure B.6. Also, the mean from the NLSCY nationally representative sample is represented by the solid line running across Figure B.6. The interaction

between site (Better Beginnings vs. comparison) and time (1998 vs. 2003) was significant for both the Sudbury ($p < .05$; $E.S. = .27$) and Cornwall ($p < .01$; $E.S. = .42$) Better Beginnings sites. This resulted from the substantial increase in teacher ratings of hyperactivity-inattention from 1998 to 2003 in the Ottawa-Vanier comparison site, while the Sudbury and Cornwall Better Beginnings sites showed slight decreases over the same time period. The interaction between the combined Better Beginnings vs. combined comparison sites from 1998 to 2003 was also significant ($p < .05$; $E.S. = .21$).



Overall, these results indicate that teachers rated Grade 3 children in Better Beginnings sites lower in hyperactive-inattentive behaviours than teachers in the comparison sites and this difference was greater in 2003 than it had been in 1998, resulting primarily from the increased ratings of hyperactive-inattentive behaviours in 2003 by teachers of Grade 3 children in the Ottawa-Vanier comparison site.

Emotional-anxiety disorder. The site averages (means) of the teacher ratings of these problems are presented in Figure B.7, along with the combined means for the three older child Better Beginnings sites and for the two comparison sites. Also, the mean from the NLSCY nationally representative sample is represented by the solid line running across Figure B.7.

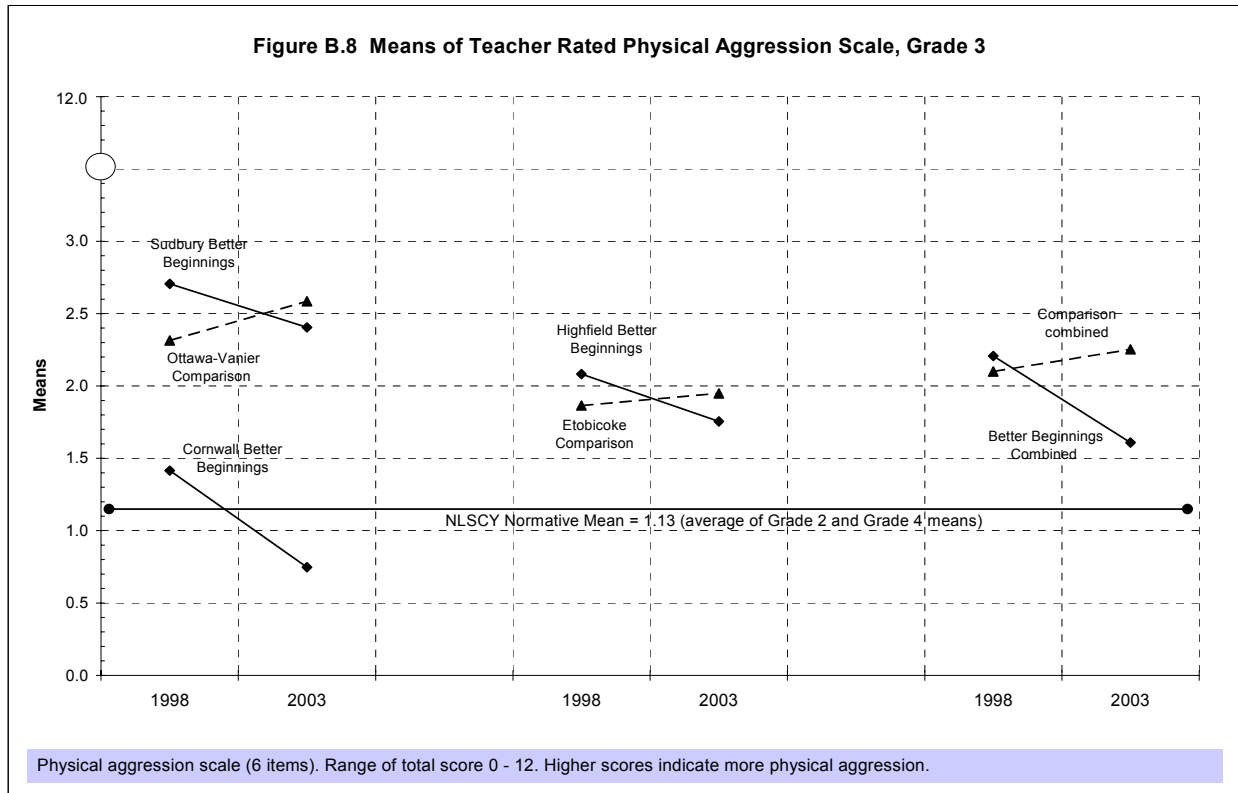


None of the interactions reflecting differences between a Better Beginnings site and its comparison site in 1998 vs. 2003 was statistically significant. Nor was the interaction between the combined Better Beginnings vs. comparison sites and the two time periods significant.

As shown in Figure B.7, the relatively high teacher ratings in Sudbury and relatively low mean ratings in Cornwall were quite stable from 1998 to 2003, while ratings by teachers in the Ottawa-Vanier comparison site showed a substantial increase from 1998 to 2003. The mean ratings in the Highfield Better Beginnings site remained stable from 1998 to 2003, while the ratings in their comparison site in Etobicoke showed a small decrease.

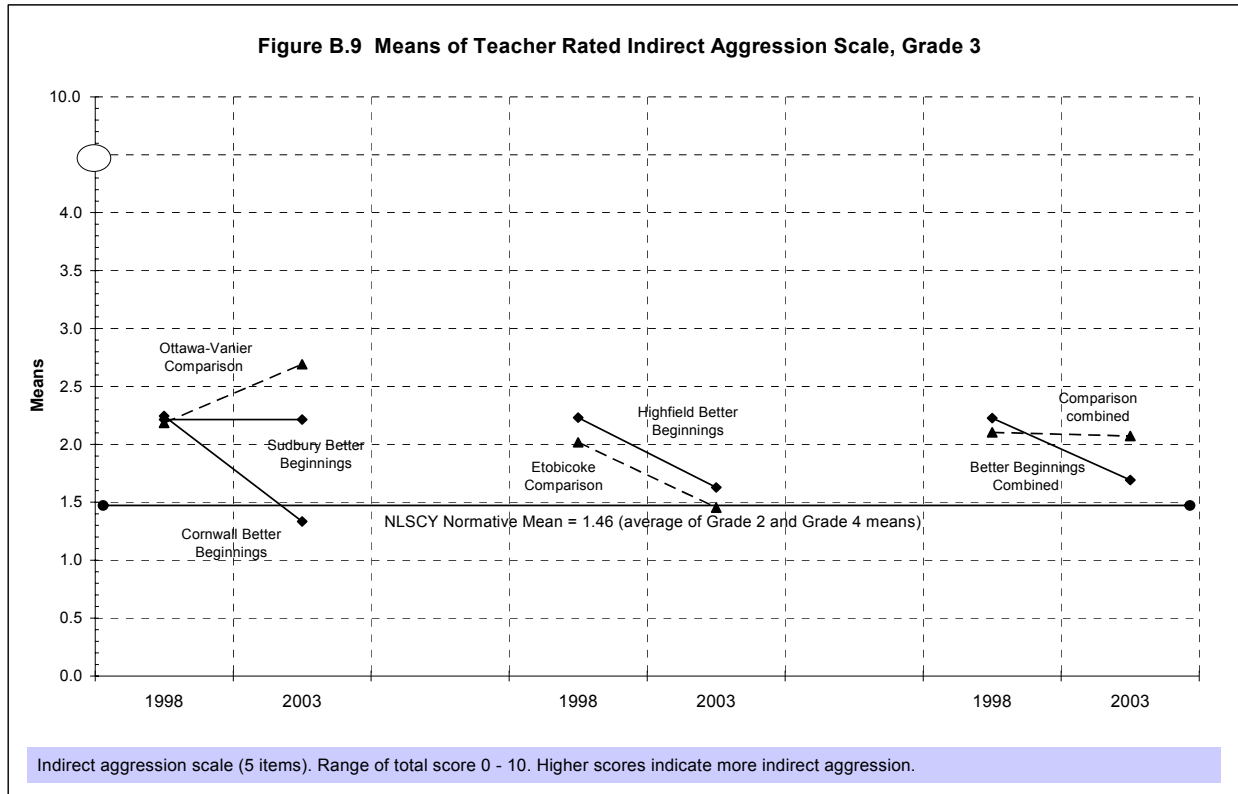
These results indicate that the lower teacher ratings of emotional problems in Cornwall and Highfield Better Beginnings sites in 1998 were maintained in 2003 relative to their comparison site, and both remained lower than the NLSCY national average. However, the Ottawa-Vanier site means increased from 1998 to 2003, the latter year ratings reaching the level of Sudbury, with both above the NLSCY national average.

Physical aggression. The mean teacher ratings of Grade 3 children on the Physical Aggression scale are presented in Figure B.8, as well as the mean from the NLSCY nationally representative sample. Several trends in these data are apparent. All three Better Beginnings sites show a decrease in teacher ratings from 1998 to 2003, while both comparison sites show an increase. Although none of the individual Better Beginnings vs. comparison site interactions are significant, the interaction effect for the combined Better Beginnings vs. comparison sites is significant ($p < .05$; $E.S. = .26$), reflecting the overall trend in each of the sites noted above.



These results indicate that although Grade 3 children’s physical aggression was rated by teachers as higher than the NLSCY normative mean, teachers in all three Better Beginnings sites rated children as decreasing in physical aggression from 1998 to 2003, unlike teachers in the comparison sites.

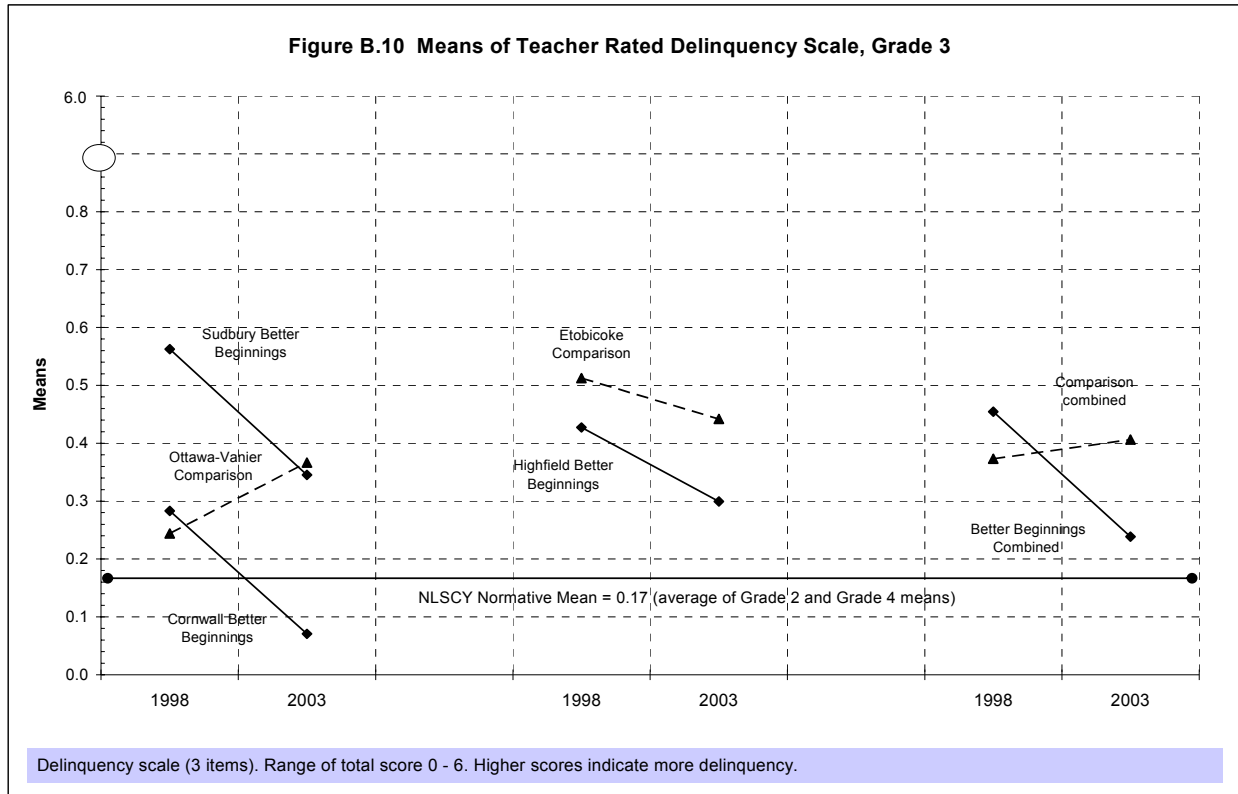
Indirect aggression. Mean teacher ratings of Grade 3 children on the indirect aggression scale appear in Figure B.9. Also, the mean from the NLSCY nationally representative sample is represented by the solid line running across Figure B.9.



The decrease in the Cornwall ratings from 1998 to 2003 contrasted with the increase over this period in the Ottawa-Vanier comparison site yielded the only significant interaction effect ($p < .05$; $E.S. = .46$). While teacher ratings in Sudbury remained stable from 1998 to 2003, ratings in both the Highfield Better Beginnings site and its Etobicoke comparison site showed equivalent decreases over the same time period. The apparent interaction between the combined Better Beginnings vs. combined comparison sites over the two time periods did not reach statistical significance.

By 2003, teacher ratings in both the Cornwall and Highfield Better Beginnings sites, as well as the Etobicoke comparison site, were at or near the NLSCY norm, while both Sudbury and Ottawa-Vanier remained substantially higher.

Delinquency. At Grade 3, delinquent behaviours, as rated by teachers, are extremely rare, averaging .17 out of a maximum score of 6.0. Somewhat higher scores begin to appear in Grade 6, but it is not until Grade 9 and beyond that these behaviours become frequent enough to be of a major social concern. However, even as early as Grade 3, site differences in teacher ratings have appeared and are presented in Figure B.10.



The general trends here are somewhat reminiscent of those seen in the ratings of other behaviour problems, i.e., a decrease in Cornwall accompanied by an increase in its comparison site in Ottawa-Vanier ($p < .05$; $E.S. = .37$), and both the Highfield Better Beginnings site and its Etobicoke comparison site showing comparable decreases. There was a sharp decrease in mean teacher ratings of delinquency in the Sudbury Better Beginnings site from .56 in 1998 to .35 in 2003, which was significantly different than Ottawa-Vanier ($p < .05$; $E.S. = .35$). The larger decreases in all three Better Beginnings sites from 1998 to 2003 relative to the increase in the two comparison sites combined yielded a significant interaction ($p < .01$; $E.S. = .25$). Despite this, only the Cornwall site in 2003 yielded a mean rating below the NLSCY average, mirroring the ratings of physical aggression reported earlier.

Summary of analyses of Grade 3 children’s emotional and behavioural problems.

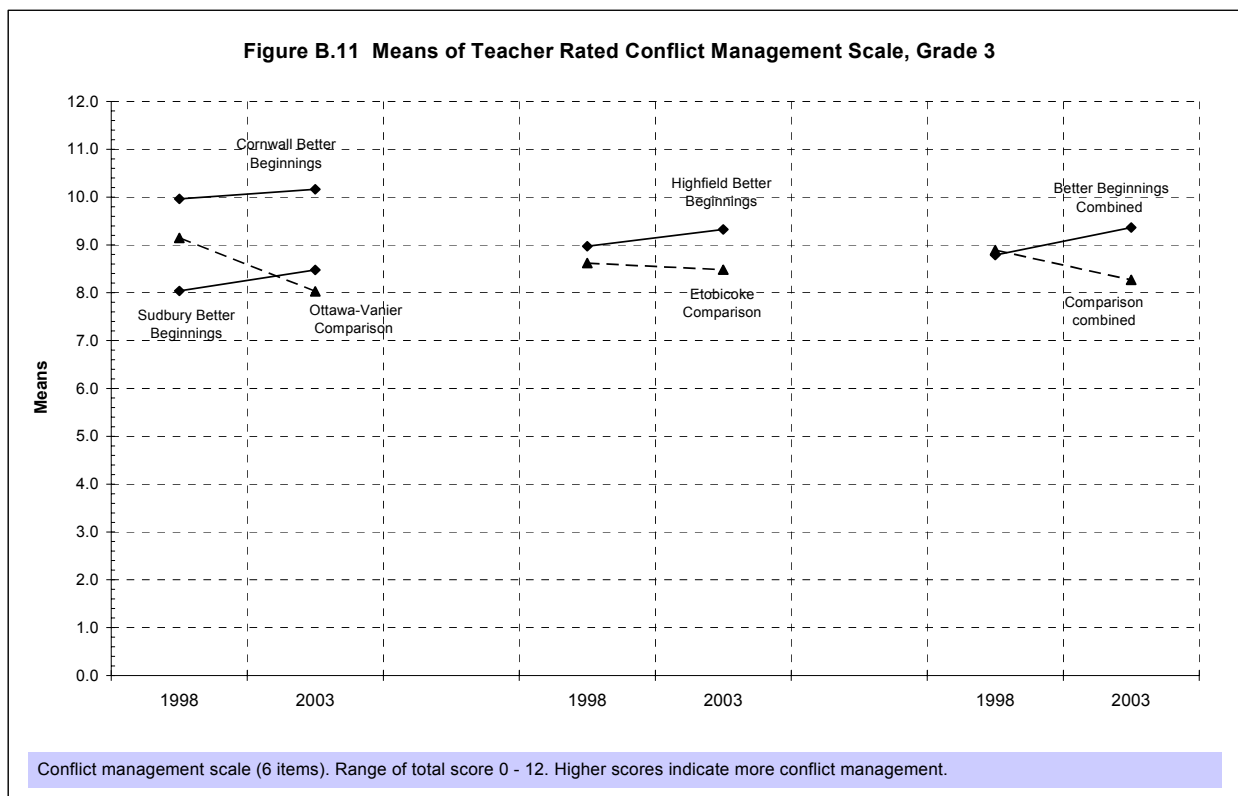
The Better Beginnings sites showed greater decreases in teacher ratings of all five measures of emotional and behavioural problems over the five year period from 1998 to 2003 than the comparison sites. The results were more consistent in the Cornwall and Sudbury Better Beginnings sites due to the finding that teacher ratings in their Ottawa-Vanier comparison site showed increases in all five measures from 1998 to 2003. Although ratings in the Highfield Better Beginnings site did not improve significantly more on any of the measures than those in its Etobicoke comparison site, the Highfield teacher ratings were lower than those from Etobicoke on four of the five measures. These outcomes indicate that the Better Beginnings programs in the three older child sites were having a positive impact on children’s behaviour in the school setting.

Results of Analyses of Grade 3 Children’s Social Skills

The development of appropriate prosocial behaviours and social skills (i.e., learning how to relate effectively with adults and other children) is one of the most critical developmental tasks for preschool and primary school children. Developing prosocial behaviour is also an important protective factor in preventing behavioural and emotional problems in early school-aged children (Eisenberg & Fabes, 1998). Consequently, fostering the development of social skills and prosocial behaviour in young children has been a key objective of Better Beginnings.

Teachers rated the Grade 3 children on the three subscales from the Social Skills Rating System (SSRS), Gresham & Elliott, 1990: conflict management, cooperation and assertiveness. As previously summarized in Table B.5, all of these Better Beginnings sites showed positive outcomes on the conflict management and cooperation measures. These results are described in more detail in the following sections.

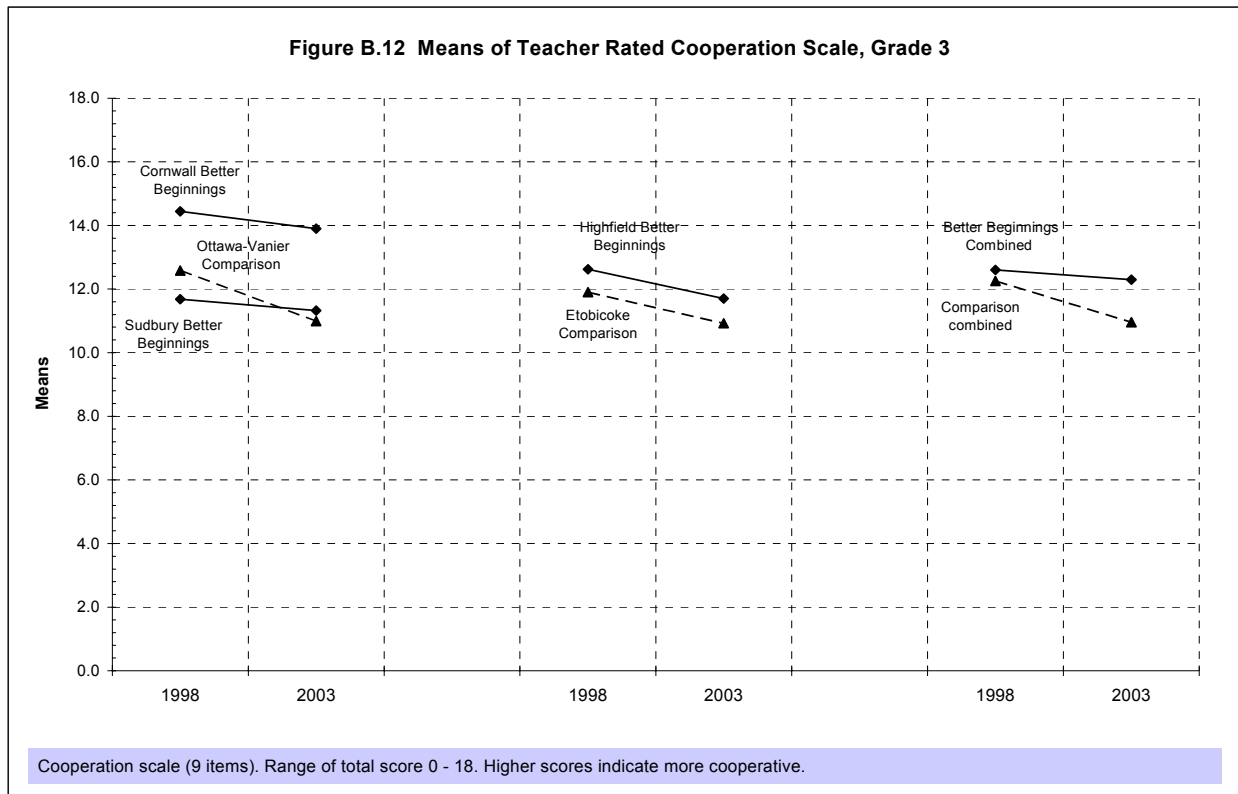
Conflict management. As depicted in Figure B.11, this 6-item subscale from the SSRS yielded small increases in mean teacher ratings from 1998 to 2003 in all three Better Beginnings sites, and a decrease in both comparison sites. The site by time interaction effect was significant for both Cornwall ($p < .05$; E.S.=.46) and Sudbury ($p < .01$; E.S.=.49) Better Beginnings sites relative to their Ottawa-Vanier comparison site, but not for the Highfield vs. Etobicoke interaction.



The three combined Better Beginnings sites vs. two combined comparison sites interaction effect was also significant ($p < .001$; $E.S. = .41$).

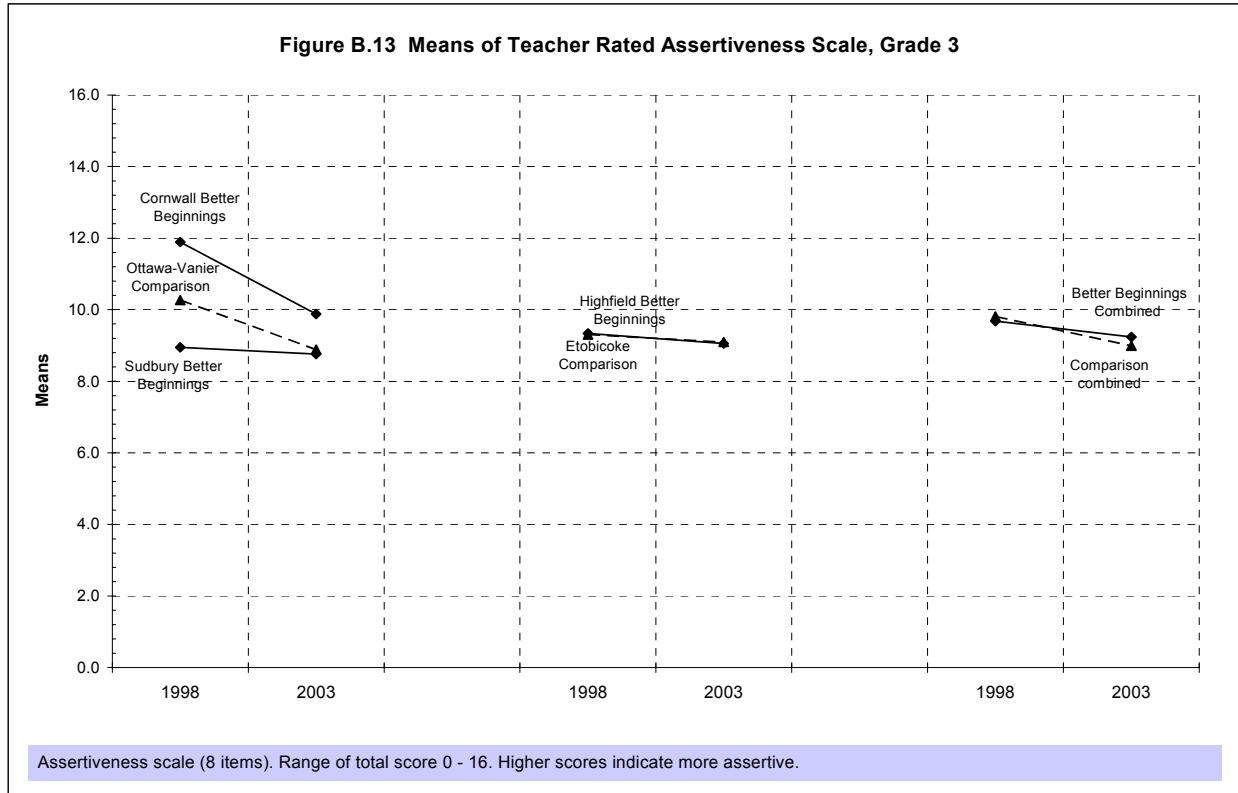
These results indicate that for this set of social skills, the Grade 3 children in the Better Beginnings sites showed higher average levels in 2003 than 1998, while comparison site children showed a decrease over the same time period. Since the SSRS subscales are not used in the NLSCY, no comparison with Canadian norms can be made on these three subscales. However, ratings on the conflict management scale were generally high, averaging approximately 9.0 out of a maximum score of 12.0.

Cooperation. For this 9-item SSRS scale, mean teacher ratings (presented in Figure B.12) decreased from 1998 to 2003 in all five sites. However, the decrease over this time period in the three Better Beginnings sites combined was significantly less than the decrease for the two comparison sites combined ($p < .05$; $E.S. = .23$), reflecting the higher levels of cooperative behaviour in the Grade 3 children in the Better Beginnings sites than in the comparison sites in 2003. Again, teacher ratings were generally high on this scale, averaging approximately 12.0 out of a maximum of 18.0.



Assertiveness. The mean teacher ratings for this 8-item SSRS subscale are presented in Figure B.13. Scores in Cornwall and Ottawa-Vanier decreased from 1998 to 2003 and remained essentially the same over this period for the other three sites, Sudbury, Highfield and Etobicoke. There were no significant interaction effects between any Better Beginnings site relative to its

comparison site over the two time periods. In 2003, all sites showed essentially the same average teacher ratings, approximately 9.0 out of 18.0, except for Cornwall where the ratings were somewhat higher at 10.0.

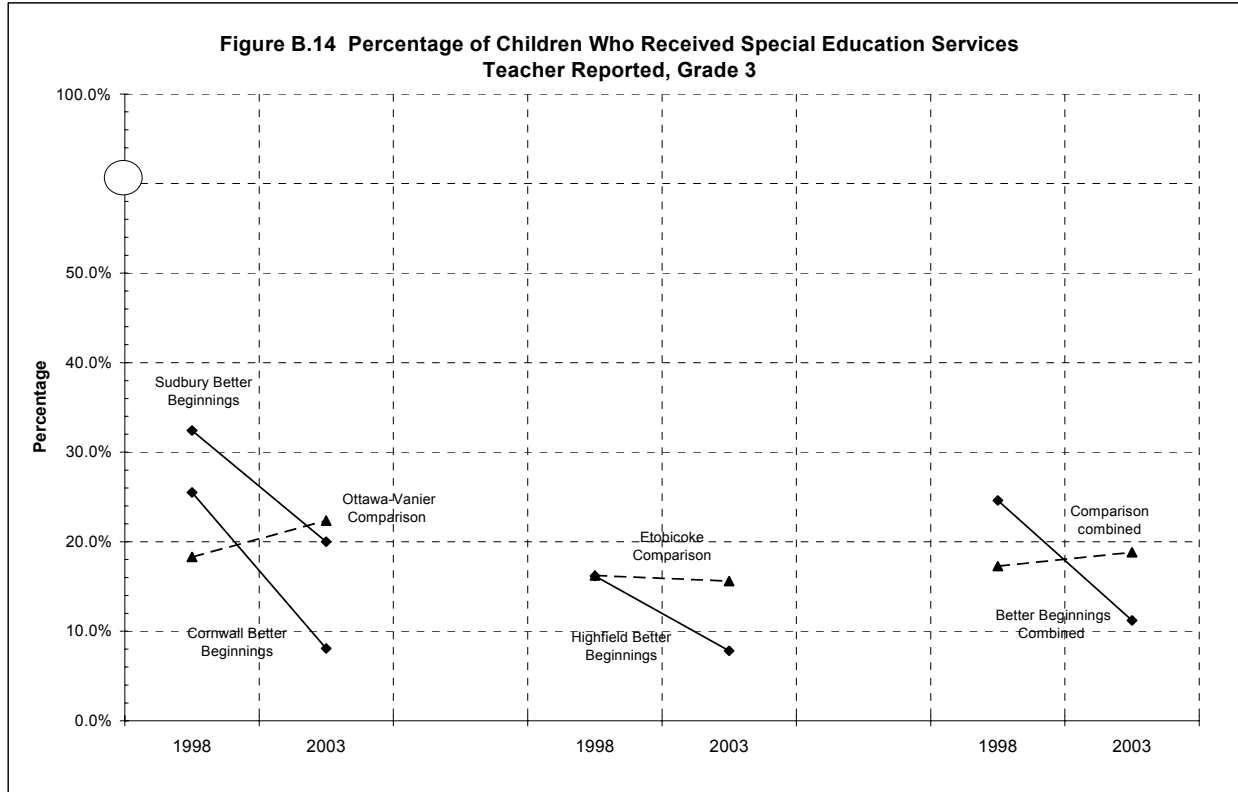


Summary of analyses of Grade 3 children’s social skills. On two of the three Social Skills Rating Scale subscales, conflict management and cooperation, teachers in the Better Beginnings sites rated their Grade 3 students as showing significantly higher average levels of these skills than teachers from the comparison sites in 2003, while no overall differences were noted in 1998 at the end of the Better Beginnings demonstration period. These results suggest that the Better Beginnings programs were having a greater impact on children in-school social skills after 1998 than they had during the demonstration phase from 1994 to 1998.

Results of School Performance Measures

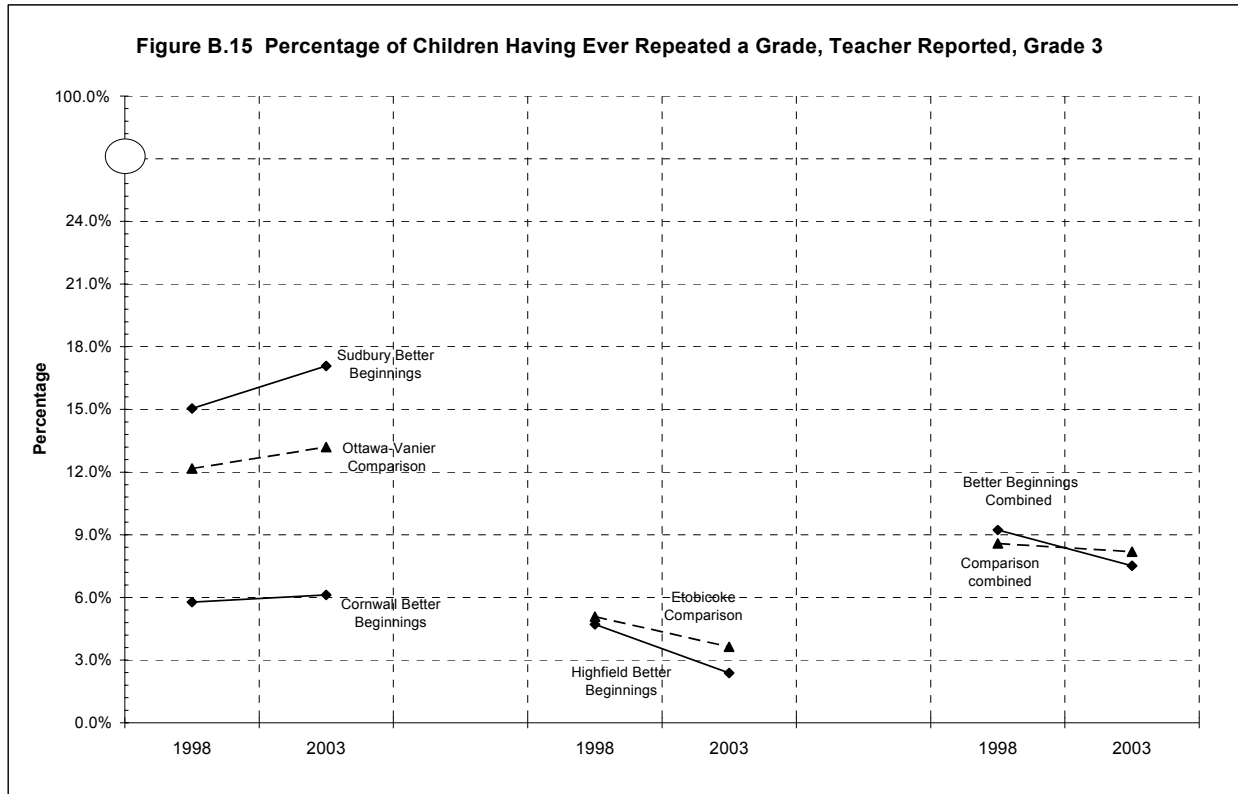
Teachers were asked to indicate whether or not the child was receiving special education services, and also whether or not the child had ever repeated a grade.

Special education involvement. The percentages of Grade 3 children reported by the teacher as receiving special education services for each site in 1998 and again 2003 are presented in Figure B.14.



The picture here is clear – the percentage of children involved in special education decreases substantially from 1998 to 2003 in all three Better Beginnings sites from 25% in 1998 to 11% in 2003, while over the same time period in the two comparison sites there is a very small increase from 17% to 19%, these changes being statistically significant for Cornwall ($p < .01$) and Sudbury ($p < .05$), but not for Highfield relative to changes in their respective comparison sites. When combined, the difference between the three Better Beginnings sites and two comparison sites was significantly different over the two time periods ($p < .001$).

Grade repetition. The percentages of children in each site who teachers reported as ever having repeated a grade are presented in Figure B.15.



The Sudbury and Ottawa-Vanier sites had substantially higher percentages of students repeating grades than the other three sites. However, there were no significant differences between Better Beginnings and comparison sites in the percentage of grade repetition in 1998 vs. 2003.

C. SPECIAL EDUCATION USE: A LONGITUDINAL NEIGHBOURHOOD-LEVEL ANALYSIS

In addition to the teacher report of the use of special education services by Grade 3 children in 1998 and 2003 reported in the previous section, data were also collated from the Principals October Report covering special education use in schools from both the Better Beginnings and comparison sites.

The Ontario Ministry of Education collects information from the principal of every school in Ontario concerning student education number, types of special education programs, and numbers of students receiving instruction for various types of special education needs. This information is provided for the school as a whole and does not distinguish one grade level from another. Thus, the Principals October Report data cover all grades in a school, usually JK to Grade 6 or Grade 8. Despite this limitation, this information is an interesting way to monitor general trends in special education services on a community level.

For this report, information on special education services from the Principals October Report from 1992 to 2001 was gathered for the schools located within each of the Better Beginnings

project sites and comparison sites. Specifically, Principals October Report data were gathered in the younger child sites as follows: 5 schools in the Guelph site, 6 schools in the Kingston site, 5 schools in the Regent Park site, 4 schools in the S.E. Ottawa site, and 18 schools in the Peterborough/Hamilton comparison site. No data were available from the Walpole Island Elementary School because it is not a provincially funded school. In the older child sites, data were gathered as follows: 6 schools in Cornwall, 1 school in Highfield, 6 schools in Sudbury, 3 schools in Etobicoke, and 13 schools in Ottawa-Vanier.

The total number of students in each school and the number of students identified by the Principal as using special education services was obtained for each school. The percentage of special education students from all the schools within each of the sites was then calculated for each year from 1992 to 2001.

ANALYSES

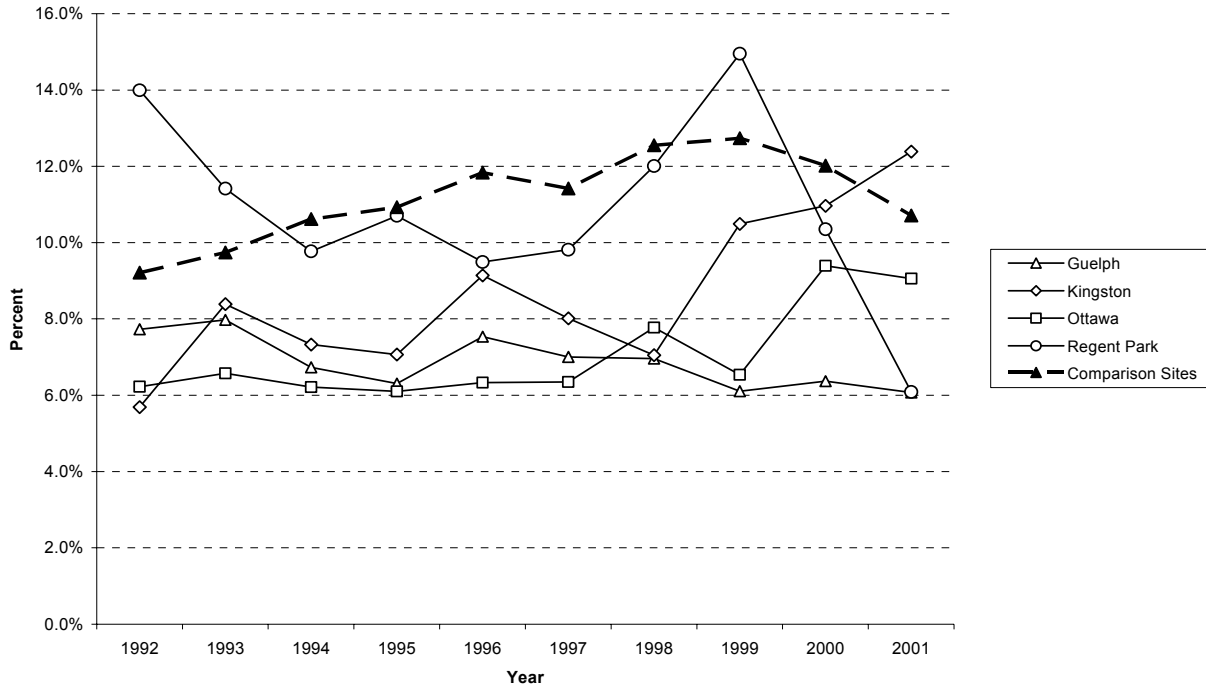
Information on special education services from the Elementary School Principals October Report from 1992 to 2001 was analyzed for the schools in each of the younger and older project sites, as well as for schools in the comparison sites. Because the dependent variable is dichotomous in nature (is the student receiving special education services – yes/no), we carried out logistic regression taking year, site and their interaction as the independent variables. We then interpreted the result of the interaction that tells whether or not the probability of students requiring special education services improved over the years in the Better Beginnings site as opposed to the comparison sites.

RESULTS

Younger Child Sites

The percentage of students receiving special education services in each of the four younger child Better Beginnings sites and the comparison sites from 1992 to 2001 are presented in Figure B.16.

Figure B.16 Percentage of Children Receiving Special Education Services, Younger Child Sites, 1992 - 2001



The logistic regression analyses of these data were designed to determine whether the changes in the percentage of special education students in each Better Beginnings site from 1992 to 2001 was significantly different from changes during this time in the comparison sites. The results of these analyses are presented in Appendix O.

These analyses indicated that the percentage of special education students for schools in the Guelph and Regent Park Better Beginnings neighbourhoods showed significantly greater decreases over the 10 year period than schools in the Peterborough and Hamilton comparison neighbourhoods ($p < .001$ for both). The percentage in Guelph started at just under 8% in 1992 and showed small but consistent decreases over the 10 year period to 6% in 2001. Percentages in the Regent Park Better Beginnings neighbourhood schools started extremely high at 14% in 1992, decreased to below 10% in 1996, increased again to a very high level of 15% in 1999, then decreased precipitously in both 2000 and 2001 to a very low level, equaling that at the Guelph schools at 6%.

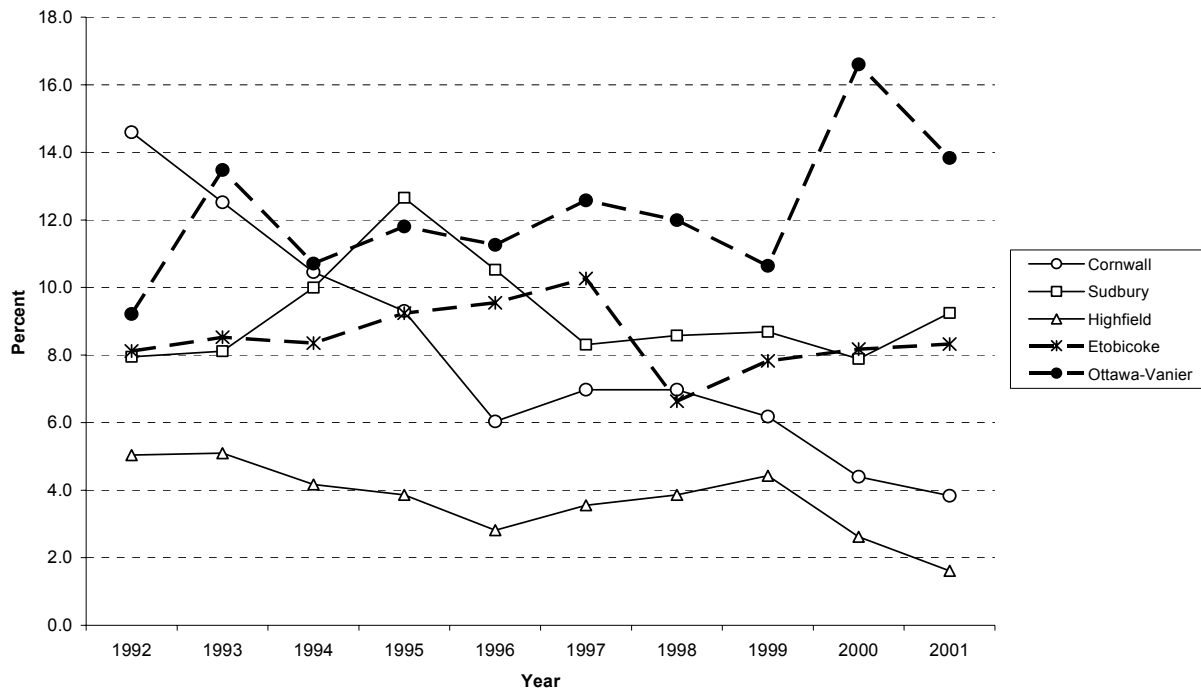
For schools in the Kingston Better Beginnings site, however, the percentage of special education students increased significantly more over the 10 year period than for the comparison sites ($p < .001$). This resulted from a marked increase in Kingston from 1998, when 7% of the students were involved in special education services, to over 12% in 2001. Although a similar rise occurred from 1999 to 2001 in the Ottawa Better Beginnings neighbourhood schools, the difference between Ottawa and the comparison sites over the 10 year period was not statistically significant.

Due to the high year to year variability in the Regent Park, Kingston and Ottawa schools, it is difficult to draw any firm conclusions from these results.

Older Child Sites

The percentages of special education students in the three older child Better Beginnings sites in Cornwall, Highfield, and Sudbury, and two comparison sites in Etobicoke and Ottawa-Vanier are presented in Figure B.17.

Figure B.17 Percentage of Students Receiving Special Education Services
Older Child Sites, 1992 - 2001



It should be remembered Ottawa-Vanier is the comparison site for Cornwall and Sudbury, and Etobicoke is the comparison site for Highfield. Results of the three logistic regression analyses are presented in Appendix P. These analyses indicate that the percentages of special education students in both the Cornwall and Sudbury Better Beginnings neighbourhood schools showed a significantly greater ($p < .001$) decrease over the 10 year period from 1992 to 2001 than the percentages in the Ottawa-Vanier comparison site schools. Although Cornwall started in 1992 with a very high percentage of over 14%, there was a steady and marked decrease over the 10 year period to 4% in 2001. During this period, the schools in the Sudbury Better Beginnings neighbourhood showed an initial increase from 8% in 1992 to over 12% in 1995. Since 1997, however, the Sudbury percentage has remained fairly constant at just above 8%. The Ottawa-Vanier comparison site schools have shown an overall increasing trend in the percentage of special education students from 1992 to 2001, resulting in the significant differences with both Cornwall and Sudbury.

The decreased percentage of special education students in the Highfield Better Beginnings neighbourhood school was marginally greater than for schools in its comparison site, Etobicoke, ($p < .06$) but remained consistently the lowest of all sites over the 10 year period.

These analyses indicate that the older child Better Beginnings schools have been showing significantly greater decreases in the percentage of special education students than schools in the comparison sites since the Better Beginnings project began. The results are consistent with those from the Grade 3 teacher ratings of the percentage of special education students in Grade 3 reported in the previous section. There were significantly greater reductions in the percentage of Grade 3 children receiving special education services between 1998 and 2003 in the Better Beginnings sites than the comparison sites.

The strong emphasis on in-school and in-classroom Better Beginnings programs in Cornwall and Highfield, as well as the before and after school Better Beginnings programs in Sudbury may be important contributors to these outcomes.

CONCLUSIONS

In this section, the analyses of teacher ratings of various aspects of children's functioning in the younger child and older child Better Beginnings and comparison neighbourhoods were reported. The results generally indicated more positive child outcomes from schools in the Better Beginnings neighbourhoods than from schools in the various comparison communities. Also, the size of these differences were generally greater in 2003 than they had been in 1998/9, immediately after the end of the project demonstration period. These findings support the view that the Better Beginnings programs not only continued to have positive impacts on children's social, emotional, behavioural and academic functioning, but that these impacts strengthened over this five year period.

There were certainly differences in the strength and nature of these outcomes among the various outcome measures, among the sites, and between the two age groups.

In the younger child sites, the positive Better Beginnings effects were strongest and most consistent on the teacher ratings of children's hyperactive-inattentive behaviours, and the two school readiness EDI domains of language / cognitive development and communication / general knowledge. Overall, the effects were generally more consistently positive and stronger in the Ottawa, Regent Park and Walpole Island Better Beginnings schools than in those from Guelph and Kingston.

In the older child sites, there were also generally positive Better Beginnings effects in teacher ratings of Grade 3 children's behaviour problems and social skills. Also, both teacher ratings and school record data indicated less use of special education resources in schools from the Better Beginnings communities. These positive Better Beginnings effects in the older child sites improved more over time in Cornwall and Sudbury than in Highfield.

Differences between sites and age groups are discussed in more detail in the following section in conjunction with the major findings from the research on the sustainability of the local project organizations.

These outcome results provide encouraging evidence concerning the sustainability of the impacts of the local Better Beginnings programs over the first five years following the end of the demonstration phase of the project in 1997.