Prevention & Treatment

Prevention & Treatment, Volume 6, Article 31, posted December 18, 2003

A Meta-Analysis of Longitudinal Research on Preschool Prevention Programs for Children

Geoffrey Nelson and Anne Westhues
Wilfrid Laurier University

Jennifer MacLeod
Wellington-Dufferin-Guelph Health Unit

ABSTRACT

The objectives of this research were to determine the effectiveness of preschool prevention programs for disadvantaged children and families in the short-term (preschool), medium-term (kindergarten–8th grade), and the long-term (high school and beyond) and to identify factors that moderate program success. Meta-analysis was used to examine the effect sizes (d) of different outcome domains of 34 preschool prevention programs that had at least one follow-up assessment when the children were in school. Although cognitive impacts resulting from these programs were greatest during the preschool period (d = 0.52), they were still evident during kindergarten through Grade 8 (K–8; d = 0.30). Social–emotional impacts on children were similar at K–8 (d = 0.27) and high school and beyond (d = 0.33), as were parent–family wellness impacts at preschool (d = 0.33) and K–8 (d = 0.30). As predicted, cognitive impacts during the preschool time period were greatest for those programs that had a direct teaching component in preschool. Also as predicted, cognitive impacts during the K–8 time period were greatest for those programs that had a follow-through educational component in elementary school. The longer the intervention for children, the greater were the impacts on preschool cognitive outcomes and child social–emotional outcomes at K–8; the more intense the intervention for children, the greater were the impacts on preschool cognitive outcomes and parent–family outcomes at K–8. The largest impacts on preschool cognitive outcomes and child social–emotional and parent–family outcomes at K–8 were found for those programs that served predominantly African American children. These results indicate that preschool prevention programs do have positive short-, medium-, and long-term impacts on several outcome domains. The findings are discussed in terms of contemporary trends in and future directions for policies and preschool prevention programs for children and families.

We thank Sean Cameron, Annette Katchaluba, Kathleen Mackey, Geoff Navara, Allison Rice Roberts, and Roslyn Shields for their assistance with this project. Thanks also to Keith Horton, who helped with the calculation of effect sizes for one of the studies and to Mike Pratt for his comments on an earlier draft of this article. This research was supported by an Interdisciplinary Research Grant from Wilfrid Laurier University.

Correspondence concerning this article should be addressed to Geoffrey Nelson, Department of Psychology, Wilfrid Laurier University, Waterloo, Ontario, Canada, N2L 3C5.

E-mail: gnelson@wlu.ca

Children whose families live in poverty are exposed to multiple disadvantages that are linked to a number of developmental problems in both the short-term and the long-term (Keating & Hertzman, 1999; McLoyd, 1998; Yoshikawa, 1994). At the same time, however, there is a growing body of evidence that child and family wellness can be promoted and a variety of childhood problems (e.g., health, social, academic) can be prevented through mental health promotion and prevention programs for disadvantaged children and families (Nelson, Prilleltensky, & Peters, 2003; Prilleltensky, Nelson, & Peirson, 2001). Moreover, it appears that prevention programs, beginning either prenatally or during the preschool years, can have lasting positive impacts on children as they grow into their adolescent and adult lives (Brooks-Gunn, 2003; Karoly et al., 1998; Nelson, Laurendeau, Chamberland, & Peirson, 2001).

In this article, we examine the short-term (preschool), medium-term (kindergarten through Grade 8), and long-term (high school and beyond) impacts of preschool prevention programs on disadvantaged children's cognitive and social development and parent–family wellness, using the technique of meta-analysis. Preschool prevention programs are designed to promote children's competence and well-being and/or to prevent negative outcomes for children. These programs can be either universal (for all children) or selective (i.e., targeted at children who are at high risk for developing problems; Nelson et al., 2003). We do not include in this review studies of *early intervention*, which refers to secondary prevention or indicated prevention for children who are already showing early signs of problems in living (Nelson et al., 2003).

A meta-analytic review can provide statistical estimates of the relative effectiveness of preschool prevention programs on multiple outcome domains, both in the short-term and the long-term, as well as test hypotheses about potential moderators of program effectiveness (Durlak & Lipsey, 1991; Lipsey & Wilson, 2001; Rosenthal, 1995). Durlak and Wells (1997, 1998) completed valuable reviews of different types of prevention programs for children. However, their reviews are now somewhat dated (they covered studies through 1991); they did not focus exclusively on preschool programs; one of their reviews (Durlak & Wells, 1998) focused on early intervention programs, not prevention programs; and their reviews included studies that did not have a follow-up assessment. More recently, MacLeod and Nelson (2000) completed a review of programs designed to prevent child maltreatment and promote family wellness. Like Durlak and Wells (1997, 1998), they did not focus exclusively on preschool programs; their review included studies of early intervention; they examined only studies that reported impacts on parent–child relationships and families, not child outcomes; and they included studies that did not have a follow-up assessment. There are now enough longitudinal evaluations of preschool prevention programs to make a meta-analytic review possible.

The framework for this meta-analytic review consisted of three components: (a) outcomes related to children's cognitive development, social–emotional behavior, and parent–family wellness (to be described in more detail in the Method section); (b) the time period of the outcome assessment (preschool, kindergarten through Grade 8, high school and beyond); and (c) potential moderators of outcomes, including program characteristics (e.g., presence of a preschool education program and follow-through educational program, comprehensiveness, length and intensity of the program), participant characteristics (e.g., ethnoracial background), and study characteristics (e.g., country in which the study was conducted, methodological quality of the study). The framework for this meta-analytic review is of value because it can guide the search for statistical estimates of the direction and magnitude of effects of preschool prevention programs on different outcome constructs at different points in time and because it can provide direction for testing hypotheses about the moderators of program effectiveness.

There have been a number of narrative reviews of the literature on the effectiveness of preschool prevention programs (e.g., Cohen & Radford, 1999; Hertzman & Wiens, 1996; Karoly et al., 1998; MacMillan, MacMillan, Offord, Griffith, & MacMillan, 1994; McCain & Mustard, 1999; McLoyd, 1998; Mrazek & Brown, 2002; Nelson et al., 2001; Ramey & Landesman Ramey, 1998; Olds & Kitzman, 1993; Weissberg & Greenberg, 1998; Yoshikawa, 1994; Zigler, Taussig, & Black, 1992). With respect to the first two components of the framework for this review (outcome domain and time period), previous reviews and individual studies provide some

suggestions about the relative stability of different effects of preschool prevention programs on different outcome domains over time. One of the conclusions of these reviews is that preschool prevention programs that include a direct teaching component for children (i.e., center-based preschool education programs like Head Start) have immediate impacts on children's cognitive development but that these gains typically fade when the children enter elementary school (Consortium for Longitudinal Studies, 1983; Hubbell et al., 1985; Lazar & Darlington, 1982).

With regard to the outcome domains of social–emotional behavior and parent–family wellness, there is evidence that preschool prevention programs have positive impacts on these outcome domains. Programs such as Head Start have produced positive long-term impacts on children's social development (e.g., reduced rates of delinquency and criminal behavior) and their life course as adults (Yoshikawa, 1994; Zigler et al., 1992). For example, the evaluation of the Perry Preschool Program, a preschool education program with a home visitation component, showed striking long-term impacts. At age 27, three times as many participants in the Perry Preschool Program as control participants earned $2,000 or more per month (29% vs. 7%) and owned their own homes (36% vs. 13%), and five times as many control participants had five or more arrests than project participants (35% vs. 7%) (Schweinhart & Weikhart, 1993).

Another example of positive long-term effects on children's social–emotional behavior and parent–family wellness can be found in the evaluation of the Prenatal/Early Infancy Project, a nurse home visitation program (Olds & Korfmacher, 1998). In the short term, positive impacts of the program on rates of child maltreatment (4% for the intervention group compared with 19% for the control group) and other outcome indicators were found for the subsample of participants who had the most risk factors (Olds, Henderson, Chamberlin, & Tatelbaum, 1986). Over time, the positive impacts of the intervention became more pronounced. In a 15-year follow-up study, Olds et al. (1997) found that for the subsample of women with the most risk factors, those who were visited by nurses had higher rates of employment than the women in the control group, as well as lower rates of impairments due to alcohol or substance abuse (41% vs. 73%), verified child abuse or neglect (29% vs. 54%), arrests (16% vs. 90% according to state records!), convictions, days in jail, use of welfare, and subsequent pregnancies. Also, when the children were 15 years of age, compared with those whose mothers were in the control group, those whose mothers participated in the home visitation program had significantly fewer incidents of running away (24% vs. 60%), arrests (20% vs. 45%), and convictions and violations of probation (9% vs. 47%) (Olds et al., 1998).

With regard to the third component of the framework for this review, there are three types of moderators of program effectiveness on the different outcome domains at the different time periods: (a) program characteristics, (b) participant characteristics, and (c) study characteristics. Several program characteristics have been identified as potentially important moderators of outcomes. Reviewers have suggested that preschool intervention programs that provide direct educational activities for children in a preschool center are more likely to promote cognitive development than programs that do not have an educational component for children (Cohen & Radford, 1999; Ramey & Landesman Ramey, 1998). Programs that focus on children only indirectly through a parent or family support intervention, such as home visitation programs, are far less likely to have an impact on cognitive outcomes than programs with a direct teaching component for children. Moreover, there is some evidence suggesting that follow-through programs that continue to provide academic assistance to children when they enter elementary school help to maintain the gains from preschool education programs (McLoyd, 1998; Ramey & Landesman Ramey, 1998).

Another potentially important program characteristic that may moderate effectiveness is the comprehensiveness of the program. While there are some programs that focus exclusively on children (e.g., the early Head Start programs) and some programs that focus primarily on parents or families (e.g., parent training or home visitation programs), programs that are more multidimensional and offer several different program components (programs that focus on children and programs that focus on parents) may achieve better outcomes than more narrowly focused programs. An early example of a comprehensive program is the Parent Child Development Centers (Andrews et al., 1982). These programs consisted of

> a comprehensive curriculum for mothers, consisting of information on child development and child-rearing practices, home management, nutrition and health, mothers' personal development,

> and government and community resources; a simultaneous program for their children; and extensive support services for participating families. (Andrews et al., 1982, p. 1)

Previous reviews of the literature have suggested that comprehensive, multicomponent programs that provide both preschool education and home-based support have positive outcomes on children's cognitive and social–emotional behavior and parent–family wellness (Nation et al., 2003; Nelson et al., 2001; Weissberg & Greenberg, 1998; Yoshikawa, 1994; Zigler et al., 1992). In their meta-analysis, MacLeod and Nelson (2000) found the highest effect sizes (d = 0.58) for comprehensive, multicomponent programs compared with single-component programs on measures of parent–family wellness.

Narrative reviews have also suggested that the length and intensity of the intervention may be important moderators of program impacts on children and parents (Cohen & Radford, 1999; McLoyd, 1998; Nation et al., 2003; Nelson et al., 2001; Ramey & Landesman Ramey, 1998; Weissberg & Greenberg, 1998; Yoshikawa, 1994). In their meta-analytic review of the impact of home visitation programs on child maltreatment and family wellness, MacLeod and Nelson (2000) found that programs had to be longer than 6 months in duration and have more than 12 visits to be effective. Reviewers have also argued that programs that begin early in the child's life (i.e., at the time of birth or prenatally) may be more effective than programs that begin later in the preschool years (e.g., ages 3 or 4, when center-based programs like Head Start typically begin) (e.g., Cohen & Radford, 1999; Nelson et al., 2001; Ramey & Landesman Ramey, 1998; Weissberg & Greenberg, 1998; Yoshikawa, 1994).

A second group of moderator variables consists of characteristics of the participants in the interventions. Ethnoracial background and socioeconomic status (SES) are potential moderators of program success. Lee, Brooks-Gunn, and Schnur (1988) found greater short-term impacts of Head Start on African American children than on White European American children. In her review of this study, McLoyd (1998) explained this finding in the following way:

> In general, children who started out the lowest gained the most. African American children were more likely to be big gainers than White children because they were relatively more disadvantaged demographically and scored significantly lower on all four measures in the preintervention year (Lee et al., 1988). (p. 163)

African American children may benefit more from preschool prevention programs because of the severe economic disadvantages and related environmental stressors (high rates of neighborhood poverty, crime, and violence) that they face (McLoyd, 1998). Not only are African American children more likely to live in poverty than White European American children, but they live in the most dire and prolonged states of poverty (McLoyd, 1998).

The last group of moderator variables consists of study characteristics. Where the study was conducted, whether it was published or not, and the methodological quality of the study are potential moderators of program effectiveness. Durlak (1995) suggested that when methodological criteria are not used to exclude studies from a meta-analysis, the meta-analyst must either demonstrate that methodological factors do not account for the variance in effect sizes or control for such factors that are found to be related to effect sizes.

In summary, the purpose of this study was to conduct a meta-analytic review of the effectiveness of preschool prevention programs. With respect to the first two components of the framework for the review (outcome domains and time periods), we descriptively examined the average effect sizes for each of the outcome domains at each of the time periods for which a sufficient number of studies was available. With regard to the third component of the framework (moderators of program effectiveness), we tested the following hypotheses based on our reading of previous narrative reviews:

> Those programs that have a direct teaching component (i.e., a preschool education program) will have significantly greater impacts on cognitive outcomes than those programs that do not have a direct teaching component.

Cognitive outcomes during the school years will be significantly stronger for programs that have a follow-through educational program into elementary school than for programs without a follow-through program.

Program qualities, such as program comprehensiveness, length, and intensity, will be significantly directly related to outcome effect sizes.

Other potential moderators of effectiveness, including participant and study characteristics were also examined.

Method

Literature Search

The following inclusion and exclusion criteria were used for the selection of studies for the review:

1. the intervention must have begun during a child's preschool years (i.e., before kindergarten or age 5, which is when children typically start school in the United States);
2. the intervention focused on the promotion of child, parent–child, or family well-being (i.e., it was a prevention or promotion program, either universal or selective, for children who did not manifest mental health or developmental problems);
3. research was reported in journal articles, book chapters, books, unpublished reports, and dissertations through the year 2000;
4. a prospective design with a control or comparison group was used;
5. outcome measures that were indicators of the constructs of children's cognitive development, children's social–emotional behavior, or parent–family wellness were collected and reported;
6. studies had to include at least one follow-up assessment into elementary school or beyond on at least one of the three outcome constructs noted in point 5; and
7. studies were reported in a format from which effect sizes could be calculated.

All studies that met these criteria were included. A number of well-known programs such as Parents as Teachers and Healthy Families America were not included in the review because they did not meet one or more of the criteria, often the requirement that there be a follow-up into the school years. As is common practice in meta-analysis, studies that focused on different risk factors were included in the analysis. For example, in the large meta-analyses of primary and secondary prevention programs for children published by Durlak and Wells (1997, 1998), they included studies that focused on a wide range of risk factors and outcome dimensions.

The literature search was conducted through a manual search of issues of journals, published between 1993 and 2000, that carry preschool prevention studies (*American Journal of Community Psychology*, *Canadian Journal of Community Mental Health*, *Child Abuse and Neglect, Child Development, Child Welfare, Journal of Community Psychology, Journal of Consulting and Clinical Psychology*, *Journal of Primary Prevention*, *Pediatrics*, and *Prevention in Human Service* (now *Journal of Prevention and Intervention in the Community*), an examination of relevant review articles and the reference sections of studies, asking other researchers for relevant sources, and an examination of eight databases covering the years from 1970 through 2000 (e.g., ERIC, PsycINFO, Dissertation Abstracts) using the following keywords: prevention, children, preschool education, home visitation, multicomponent program. As advised by Lipsey and Wilson (2001, p. 30), efforts were made to track down unpublished studies that were identified through the search. Although we cannot be certain that we located all the studies that met the criteria listed above, we believe that we found most of these studies. We reviewed over 5,000 abstracts and our search uncovered all of the studies that have been included in several recent narrative reviews of this literature (Cohen & Radford, 1999; Hertzman & Wiens, 1996; Karoly et al., 1998; MacMillan et al., 1994; McCain & Mustard, 1999; McLoyd, 1998; Mrazek & Brown, 2002; Nelson et al., 2001; Ramey & Landesman Ramey, 1998; Olds & Kitzman, 1993; Weissberg & Greenberg, 1998; Yoshikawa, 1994; Zigler et al., 1992).

Unit of Analysis

The unit of analysis for this research was "program intervention." In a few cases, both the intervention and comparison groups received an intervention (e.g., a preschool program). In such cases, we included only those studies in which the intervention program participants received something additional (e.g., a parenting program) or another component that was hypothesized to be something more than the participants in the comparison group received.

In studies that compared more than one intervention or in which the different programs were implemented in different sites, the participants in the intervention groups were sometimes compared with the same group of participants in the control or comparison group. This introduces the problem of statistical dependence of effect sizes across interventions (Lipsey & Wilson, 2001, p. 126). Steps were taken to eliminate this problem of statistical dependence.[1] A total of 34 interventions were found that met the inclusion criteria. The interventions, with relevant citations, effect sizes corrected for small-sample bias, sample sizes for intervention and control/comparison groups, and age of children at the longest follow-up period, are noted in Table 1.

Back to article

Table 1

Programs Included in the Meta-Analysis, Intervention and Control Comparisons, Average Effect Sizes (d) Adjusted for Sample Size, Sample Sizes for Outcome Domains at Different Time Periods, and Childï¿½s Age at Longest Follow-Up Period

	Cognitive outcomes		Socialï¿½emotional outcomes		Parentï¿½family outcomes	
Program and comparison	Preschool (*k* = 17)	Kï¿½8 (*k* = 27)	Kï¿½8 (*k* = 19)	High school and beyond (*k* = 10)	Preschool (*k* = 7)	Kï¿½8 (*k* = 12)
Better Beginnings, Better Futures (Peters et al., 2000)ï¿½multicomponent program versus comparison, nonrandom assignment to condition						
Cornwall site						
d		0.00	0.06			0.03
Experimental *n*		66	66			66
Control *n*		44	42			42
Follow-up age		8	8			8
Highfield site						
d		0.00	0.24			0.42
Experimental *n*		79	79			79
Control *n*		43	43			43
Follow-up age		8	8			8
Sudbury site						
d		0.00	–0.05			0.05
Experimental *n*		112	112			112
Control *n*		118	118			118
Follow-up age		8	8			8
Bright Start (Tzuriel et al., 1999)ï¿½multicomponent program vs. control, nonrandom assignment to condition						
d	0.41	0.63				
Experimental *n*	82	43				
Control *n*	51	36				
Follow-up age		5				
Brookline Early Education Program (Pierson et al., 1983)ï¿½BEEP vs. control, nonrandom assignment to condition						

d		0.23	0.15		
Experimental *n*		169	169		
Control *n*		169	169		
Follow-up age		5	5		

Brooklyn (Meyer, 1984)ï¿½DISTAR follow through vs. comparison, nonrandom assignment to condition

d				0.36	
Experimental *n*				65	
Control *n*				100	
Follow-up age				18	

Busselton Study (Cullen, 1976)ï¿½visits with family physician focused on parenting vs. contro, random assignment to condition

d		0.00	0.12	0.21	0.17
Experimental *n*		124	105	124	124
Control *n*		122	104	122	122
Follow-up age		6	6	29	6

Carolina Early Intervention Programï¿½Abecedarian Project (Ramey et al., 1988)ï¿½5-year intervention and 3-year follow-through vs. control, random assignment to condition

d	0.71	0.29	0.36	0.23	0.46
Experimental *n*	57	47	25	25	25
Control *n*	54	44	23	22	23
Follow-up age		12	12	15	12

Chicago Child Parent Center (Reynolds, 1994)ï¿½CPC preschool vs. comparison, nonrandom assignment to condition

d	0.46	0.34	0.47	0.04	0.64
Experimental *n*	266	266	266	956	266
Control *n*	125	125	125	306	125
Follow-up age		11	11	16	11

Comprehensive Child Development Program (St. Pierre, Layzer, Goodson, & Bernstein, 1997)ï¿½multisite, multicomponent home visitation vs. control, random assignment to condition

d	0.04	0.03	0.04	0.02	–0.06
Experimental *n*	1,443	1,384	1,384	1,443	1,384
Control *n*	1,432	1,440	1,440	1,432	1,440
Follow-up age		5	5		5

Early Training Project (Gray & Klaus, 1970)ï¿½home visitation and preschool vs. control, random assignment to condition

d	0.70	0.62		0.38	
Experimental *n*		38	38		41
Control *n*		42	42		21
Follow-up age		8		16	

Five Site (Becker & Gersten, 1982)ï¿½follow-through in 5 sites vs. preschool-only control, nonrandom assignment to condition

d		0.25
Experimental *n*		624
Control *n*		567
Follow-up age		11

HIPPYï¿½Arkansas (Baker, Piotrowski, & Brooks-Gunn, 1999)ï¿½home visitation vs. comparison, nonrandom assignment to condition

d		0.28	0.34
Experimental *n*		121	121
Control *n*		105	105
Follow-up age		6	6

HIPPYï¿½New York (Baker et al., 1999)ï¿½home visitation vs. control (both groups were in a preschool), random assignment to condition

d		0.31	0.43	
Experimental *n*		84	84	
Control *n*		98	98	
Follow-up age		7	7	

Houston Parentï¿½Child Development Center (Andrews et al., 1982)ï¿½home visitation and preschool vs. control, random assignment to condition

d	−0.01	0.17	0.31	0.28
Experimental *n*	38	51	50	41
Control *n*	53	88	87	50
Follow-up age		10	10	

Infant Health and Development (IHDSP, 1990)ï¿½multicomponent program vs. control (8 sites), random assignment to condition

d	0.62	0.34	0.05
Experimental *n*	377	336	336
Control *n*	608	538	538
Follow-up age		8	8

Institute for Developmental Studies (Deutsch, Deutsch, Jordan, & Grallo, 1983)ï¿½multicomponent program vs. comparison, nonrandom assignment to condition

d	0.35	0.54
Experimental *n*	268	135
Control *n*	172	162
Follow-up age		8

Learning to Learn (Sprigle & Schaefer, 1985)ï¿½Learning to Learn program vs. Head Start, random assignment to condition

d		0.50	0.66
Experimental *n*		44	44
Control *n*		39	39
Follow-up age		11	11

Louisville Experiment (Miller & Bizzell, 1983a, 1983b)ï¿½preschool interventions vs. control, nonrandom assignment to condition

d	0.35
Experimental *n*	114
Control *n*	36
Follow-up age	13

Motherï¿½Child Home Program (Levenstein, Levenstein, Shiminski, & Stolzberg, 1998)ï¿½home-based intervention with mothers vs. control, random assignment to condition

d	0.79	0.68	0.25
Experimental *n*	20	21	70
Control *n*	26	30	14
Follow-up age	13	13	17

New Haven (Miller & Dyer, 1975)ï¿½follow through vs. preschool only, nonrandom assignment to control

d	1.25	0.26	−0.06	0.59
Experimental *n*	61	33	33	33
Control *n*	48	26	26	24
Follow-up age		13	13	17

Optimum Growth Project (Caruso, 1989)ï¿½home visitation vs. comparison, nonrandom assignment to condition

d	0.59	0.12	0.35	0.31
Experimental *n*	171	171	66	66
Control *n*	91	91	44	44
Follow-up age		6		6

Parenting Intervention (Webster-Stratton, 1998)ï¿½parenting intervention vs. control (both groups in Head Start),

nonrandom assignment to conditiona

d		−0.01	0.36	0.36
Experimental *n*		189	264	189
Control *n*		107	130	107
Follow-up age		5		5

Perry Preschool (Schweinhart & Weikart, 1997)ï¿½Perry preschool vs. control, random assignment to condition

d	1.03	0.76	0.36
Experimental *n*	68	68	55
Control *n*	65	65	62
Follow-up age		13	18

Philadelphia Project (Beller, 1983)ï¿½nursery, kindergarten, and Grade 1 vs. Grade 1 only, nonrandom assignment to condition

d	0.50
Experimental *n*	51
Control *n*	54
Follow-up age	9

Portland and Trenton (Lee, Brooks-Gunn, & Schnur, 1988; Lee, Brooks-Gunn, Schnur, & Fong-Ruey, 1990)ï¿½13 generic Head Start programs vs. no preschool, nonrandom assignment to condition

d	0.16	0.17
Experimental *n*	333	333
Control *n*	204	204
Follow-up age	6	6

Prenatal/ Early Infancy Projectï¿½Elmira Site (Olds & Korfmacher, 1998)ï¿½home visitation vs. control, random assignment to condition

d	0.13		0.51	0.10
Experimental *n*	166		97	165
Control *n*	148		148	189
Follow-up age			15	

Project Care (Wasik, Ramey, Bryant, & Sparling, 1990)ï¿½family education and child development centers vs. control, random assignment to condition

d	1.15		0.36
Experimental *n*	14		13
Control *n*	8		22
Follow-up age			

Syracuse University Family Development Project (Honig, Lally, & Mathieson, 1982)ï¿½multicomponent program vs. comparison, nonrandom assignment to condition

d	0.36	0.41	0.45
Experimental *n*	82	51	51
Control *n*	74	42	42
Follow-up age		13	13

University of Illinois (Karnes, Schwedel, & Williams, 1983)ï¿½5 preschool interventions vs. no preschool, nonrandom assignment to condition

d	0.41
Experimental *n*	24
Control *n*	17
Follow-up age	18

Vermont Motherï¿½Infant Project (Rauh, Achenbach, Nurcombe, Howell, & Teti, 1988)ï¿½nursing intervention for mothers of low birthweight babies vs. no-treatment control, random assignment to condition

d	0.38	0.67		0.85
Experimental *n*	25	24		25
Control *n*	28	32		28

	Follow-up age		7			
Washington (Evans, 1985)ï¿½Head Start and DISTAR vs. no preschool, nonrandom assignment to condition						
	d		0.13			
	Experimental *n*		44			
	Control *n*		20			
	Follow-up age		13			
Yale Child Welfare (Seitz, Rosenbaum, & Apfel, 1985)ï¿½multicomponent program vs. comparison, nonrandom assignment to condition						
	d	0.46		0.70		0.78
	Experimental *n*	9		18		15
	Control *n*	9		17		15
	Follow-up age			10		10
Ypsilantiï¿½Carnegie Infant Education Project (Lambie, Bond, & Weikart, 1974)ï¿½home visitation vs. control, random assignment to condition						
	d	0.27	0.08			0.05
	Experimental *n*	22	20			20
	Control *n*	22	25			25
	Follow-up age		7			7
Average						
	d	0.52	0.30	0.27	0.33	0.33	0.30
	Experimental *n*	187	170	183	149	288	200
	Control *n*	178	161	177	84	271	179
	Age at longest follow-up		8.8	8.7	17.9		8.3

Note. In a few cases sample sizes at follow-up periods are larger than sample sizes at earlier time periods because researchers added additional cohorts for the follow-up research to increase the sample sizes. None of the studies reported cognitive outcomes for high school and beyond. We decided not to report socialï¿½emotional outcomes at preschool (as there were only 3 studies reporting such outcomes) or parentï¿½family outcomes at high school and beyond (as only 1 study reported such outcomes). *k* = number of studies included for each outcome dimension.

[a] Participants were not randomly assigned to conditions in the Websterï¿½Stratton (1998) study, but Head Start centers were randomly assigned to conditions.

Back to top Back to article

Coding Procedures

A total of 38 variables were coded for each study. The coded variables included program characteristics (e.g., type of program, strengths orientation, program components, length and intensity of the intervention for the parents and children, whether the program involved parents or mothers only, age of child at beginning of intervention), participant characteristics (e.g., ethnoracial background, SES), and study characteristics (e.g., country in which the study was conducted, source of publication, methodological quality). We did try to rate factors like empowerment and gender of child, but most studies did not provide enough information to code these types of variables reliably.

A total score for the methodological quality of each study was calculated based on MacMillan et al. (1994). If a study satisfied all of the criteria, it received a score of 25 (the highest possible score). Variables scored included method of sample allocation (i.e., randomized design vs. quasi-experiment); assessment of baseline comparison of intervention and comparison groups; whether inclusion and exclusion criteria were clearly defined; whether target population was clearly defined; duration of follow-up; quality of follow-up (percentage of dropouts or withdrawals); whether outcome measures were administered blind; the number of outcome measures; reliability of measures used; and the number of periods of assessment.

All of the continuous variables were recoded into categorical variables so that they could be used for the purpose of moderator variable analysis (described in a subsequent section). With the small number of studies, more of a conceptual approach to deciding cut-off levels was not feasible. We used an empirical approach to make the cut-offs, striving to have an equal or similar level of number of studies for each of the analyses.

Length and intensity of the intervention were computed separately for children and parents. Length was coded in weeks. We assumed one year of preschool or school was equal to 9 months and that there are 4.3 weeks per month. Intensity was coded in terms of intended sessions for the children or parents. Few studies reported the number of sessions that children or parents actually attended, so our estimate of intensity was the maximum number of sessions that could be attended and thus is an overestimate. A session was defined as a planned activity that was part of the intervention that lasted up to one-half day. A half-day in preschool was counted as one session, as was one home visit with a parent.

Length of intervention for the child was coded as more than or less than one year, whereas intensity was coded as more than or less than 300 sessions. Length and intensity for the child were significantly related, $\chi^2(1) = 11.6$, $p < .001$. We also computed the correlation between the continuous variables of length and intensity of the intervention for children and found a very strong association, $r = .88$, $p < .001$. Length of intervention for parents was coded as more than or less than 60 weeks, and intensity was coded as more than or less than 12 sessions. Length and intensity for the parent were significantly related, $\chi^2(1) = 11.8$, $p < .001$. The correlation between the continuous variables of length and intensity of the intervention for parents was statistically significant, $r = .49$, $p < .01$. Finally, the length and intensity of the intervention for children were not related to the length and intensity of the intervention for parents, using either chi-square analysis on the categorical variables or correlation coefficients for the continuous variables.

Coding of the studies was conducted in several steps. First, a coding system and definitions for each code were developed. Next, we divided the 34 studies among themselves and coded all of the variables for these studies. Because we examined longitudinal research, there were typically two or more articles for each intervention. Thus, we used multiple sources for coding each intervention.

Intercoder Reliability

To assess the reliability of coding procedures, we trained a research assistant in the coding definitions and procedures; the assistant then independently coded 13 of the 34 interventions. Reliability coefficients could not be calculated for two of the variables (program category and source of publication) because one of the variables was a constant (i.e., only one category was used for the coding). Seven variables (SES, type of helpers, strengths orientation, child care component, education or work component, basic needs component, children's group) were dropped because the reliabilities for these variables were lower than .65. Insufficient description of these characteristics in the articles was the major problem with the codes for which the standard of a .65 kappa for reliability was not attained. For the remaining variables, Cohen's kappa (the percentage agreement corrected for chance) (Cohen, 1960) was .85 on average (range: .65–1.00). The average kappas across codes for the agreement between the research assistant who did the reliability check and each of the three primary coders (the authors) were .76, .78, and .79.

Outcome Constructs by Time Period

We used three broad outcome constructs for the analysis. Children's cognitive development was assessed by measures of IQ, achievement tests, grades, and teacher ratings of children's academic skill or performance. Children's social–emotional behavior was tapped through parent and teacher ratings of behavior problems and social skills, self-report of self-esteem, grade retention and placement in special classes, and employment, education, and criminal behavior in adolescence. Parent–family wellness was assessed through indicators of child maltreatment (direct and proxy measures), parent–child relationships (self-report and observation), family functioning (self-report), and parent mental health, employment and/or education, and social support. Originally, we started with more fine-grained outcome constructs, but we ended up using this smaller set of broad outcome constructs because there were not enough data across the interventions to meaningfully analyze some of the

more fine-grained constructs. We decided to use grade retention and placement in special classes as indicators of social–emotional adaptation because these indicators are subject to a range of different influences and are not solely based on children's cognitive abilities. Moreover, grade retention and special class placement are associated with a number of negative social–emotional outcomes for children and adolescents. We believe that the three broad outcome categories, though not perfect, are conceptually distinct and meaningful.

The outcome constructs were examined for three time periods (preschool, K–8, and high school and beyond). Data were pooled across the preschool years to examine the immediate or short-term effects of the interventions. We originally used K–3, 4–8, and high school and beyond for the later time periods to examine medium-term and long-term effects. However, it was necessary to collapse the K–3 and 4–8 categories because there were not enough cases to permit meaningful analysis of these two time periods. As can be seen in Table 1, the average age of the children at the longest K–8 follow-up period was 8.8 years for cognitive outcomes, 8.7 years for social–emotional outcomes, and 8.3 years for parent–family outcomes, whereas the average age of the children at the longest high school and beyond follow-up period was 17.9 years.

Calculation of Effect Sizes

Effect sizes (*d*s) were calculated by subtracting the posttest (or follow-up) mean of the comparison group from the posttest (or follow-up) mean of the intervention group and dividing the result by the pooled standard deviation. If means and standard deviations were not reported, effect sizes were calculated from test statistics (typically *t* tests, chi-square tests for 2 × 2 contingency tables) using formulas outlined by Wolf (1986, p. 35). Effect sizes that were reported in the articles, book chapters, and reports were used. Effect sizes were not computed for subgroups within an intervention unless information for the total sample was not available. In this case, the effect sizes for the subgroups were pooled (e.g., effect sizes for boys and girls were averaged). A total of 721 effect sizes were examined for the interventions. The mean number of effect sizes calculated per study was 21.2 (range: 3–54).

Within each study, effect sizes were pooled for each of the three outcome constructs for each of the three time periods. That is, when there was more than one effect size per construct for a given time period, each individual effect size was calculated and averaged to yield one effect size. We believed that averaging individual effect sizes would provide a more accurate estimate of the effect size for each outcome construct at each time period than would randomly selecting one specific effect size. Findings that were reported as not significant (with no accompanying numbers) were translated into effect sizes of 0, which is the recommended conservative approach employed in other meta-analyses (Durlak & Wells, 1997, 1998; Rosenthal, 1995).

Statistical Analyses

Because effect sizes have been found to be upwardly biased for studies with small sample sizes (practically speaking, those with less than 30 per group) (Durlak, 1995), we first corrected the *d*s for small-sample bias for the different outcome constructs at the three different time periods, using the formula provided by Lipsey and Wilson (2001, p. 49, formula 3.22). Because the sample sizes varied by time period (and sometimes by outcome measure), it was necessary to correct the *d*s for the different time periods and the different constructs by different sample sizes (see Table 1). All of the *d*s that we report in this article have been corrected for small-sample bias.

Next, we tested the homogeneity of the different outcome constructs to determine whether the variance among the group of studies was due to sampling error or could be attributed to systematic differences between the studies. The homogeneity of the effect sizes was tested using the *Q* statistical test developed by Hedges and Olkin (1985). If Q_t (the overall or total *Q* statistic) is statistically significant, then the null hypothesis of homogeneity is rejected. The next step in model testing following a significant Q_t was to use moderator variables (variables that were identified a priori as possibly having an impact on the effectiveness of interventions) to attempt to divide the group of studies into homogeneous subgroups.

To test models using moderator variables, two further types of Q statistic must be calculated: Q_w (the within-group fit statistic) and Q_b (the between-group fit statistic). This is a meta-analytic analogue to the analysis of variance (ANOVA). Lipsey and Wilson (2001) suggested that a random effects model, rather than a fixed effects model, be fitted to examine moderators of effect size variability, because the observed variance is due, at least in part, to unidentified sources of random differences among studies. The random effects model decomposes observed variability of effect sizes into two components, participant-level sampling error and error attributed to other sources of variability assumed to be randomly distributed. The weighted d in the random effects model is calculated differently than that for the fixed effects model, to take these two sources of error into account (the inverse variance weight has two parameters; see Lipsey and Wilson, 2001, pp. 134–135). In keeping with the recommendations of Lipsey and Wilson (2001), we used a random effects model to test for moderators when Q_t for an outcome construct was significant and the sampling error accounted for less than 75% of the variance. For all model testing and moderator analyses reported in this article, the corrected effect sizes have been weighted by an inverse variance weight. Confidence intervals were also calculated for weighted ds. For a moderator to be significant, the Q_b must be statistically significant and the Q_w for the subgroups must not be significant (the within-group variance should be homogeneous). When this occurs, the confidence intervals for the groups that are compared do not overlap (Durlak, 1995).

Because of the statistical dependence of the data for different time periods, model testing could not be performed to examine differences on the same constructs over time, as could be done with a repeated measures or mixed effects ANOVA. We used the statistical program developed by Schwarzer (1996) to compute all of the Q statistics. This program permits an examination of random effects models, but it does not allow for analyses using continuous variables as moderators. It is for this reason that all the moderator analyses were conducted using categorical variables.

Results

Descriptive Information

In Table 2, we provide descriptive information on program, participant, and study characteristics. The most frequently provided program components were home visitation (71%), parent training (68%), and preschool education (68%), with 79% of the programs offering three or more program components. Whereas length of program was somewhat similar for children (65% of programs lasted longer than one year) and adults (47% of programs lasted longer than 60 months), the intensity of programs for children (56% of programs provided more than 300 sessions) was much greater than that for parents (50% of programs provided more than 12 sessions). In terms of participant characteristics, more than half of the interventions (56%) served predominantly African American children and families. With regard to study characteristics, less than half of the studies used a randomized design (41%) and had a sample retention rate of more than 80% at the first follow-up (47%), and only half (50%) reported blind assessment on the outcome measures. On other methodological criteria, the majority of studies fared more positively. The overall methodology scores ranged from 12 to 25 (M = 18.8, SD = 4.0).

Back to article

Table 2
Program, Participant, and Study Characteristics of the 34 Studies

Program characteristics	*k* (%)
Gender of parents	
Mothers only	12 (35%)
Both parents	22 (65%)
Home visiting (program staff visited parents in their	

homes)	
No	10 (29%)
Yes	24 (71%)
Parent training (program staff teach parents child management or educational strategies to use with their children)	
No	11 (32%)
Yes	23 (68%)
Social support (peer support from other parents)	
No	18 (53%)
Yes	16 (47%)
Preschool (teachers and program staff use educational strategies with children in a preschool center)	
No	11 (32%)
Yes	23 (68%)
Family planning (program staff teaches parents about birth control)	
No	30 (88%)
Yes	4 (12%)
Parentï¿½child interaction (program staff provide program activities to promote parent-child interaction, e.g., reading, homework)	
No	16 (47%)
Yes	18 (53%)
Parentsï¿½ group (any group activity exclusively for parents, e.g., support group, group parent training)	
No	19 (56%)
Yes	15 (44%)
Total number of components	
1 or 2	7 (21%)
3 or more	17 (79%)
Length of interventionï¿½child	
52 weeks or less	12 (35%)
More than 52 weeks	22 (65%)
Length of interventionï¿½parents	
60 weeks or less	18 (53%)
More than 60 weeks	16 (47%)
Intensity of interventionï¿½child	
300 sessions or fewer	15 (44%)
More than 300 sessions	19 (56%)
Intensity of interventionï¿½parent	
12 sessions or fewer	17 (50%)
More than 12 sessions	17 (50%)
Age of child at start	
Birth to 3	23 (68%)
4 or more	11 (32%)

Participant characteristics	
Ethnoracial background	

Predominantly African American (> 50% of participants are African American)	19 (56%)
Other (< 50% of participants are African American)	15 (44%)
Study characteristics	
Randomized design	
No	20 (59%)
Yes	14 (41%)
Demonstrated similarity of intervention and comparison groups at baseline (data provided on comparability of groups or covariate adjustment performed)	
No	8 (31%)
Yes	26 (69%)
Inclusion/exclusion criteria clearly defined	
No	1 (3%)
Yes	33 (97%)
Target population clearly defined	
No	1 (3%)
Yes	33 (97%)
Follow-up duration 2 years or more	
No	4 (12%)
Yes	30 (88%)
Follow-up quality (80% or more participants retained at first follow-up)	
No	17 (50%)
Yes	16 (47%)
Missing	1 (3%)
Majority of outcome measures done blind	
No	17 (50%)
Yes	17 (50%)
2 or more outcome measures	
No	1 (3%)
Yes	33 (97%)
Reliability of outcome measures described or evident	
No	6 (18%)
Yes	28 (72%)
Total methodology score	
19 or less (low)	13 (38%)
20 or more (high)	21 (62%)
Year published	
Before 1980	13 (38%)
1980 or later	21 (62%)

Back to top Back to article

In Figure 1, we present the mean *d*s, corrected for small sample size, for the outcome constructs at each of the three time periods. There were no studies reporting cognitive outcomes at high school and beyond ($k = 0$), and we decided not to report a *d* for two other time periods because of the small number of studies on which the *d*s

were based: preschool social–emotional outcomes (k = 3) and parent–family outcomes for high school and beyond (k = 1). In meta-analysis, ds at a 0.2 level are considered small; medium ds are at a 0.5 level; and large ds are at a 0.8 level (Lipsey & Wilson, 2001). The preschool cognitive outcomes were at a medium level, and all of the other outcomes fell between small and medium effect sizes. As the reader can see from Figure 1, the average corrected ds for social–emotional outcomes during K–8 and high school and beyond are quite similar, as are the average corrected ds for parent–family outcomes during preschool and K–8.

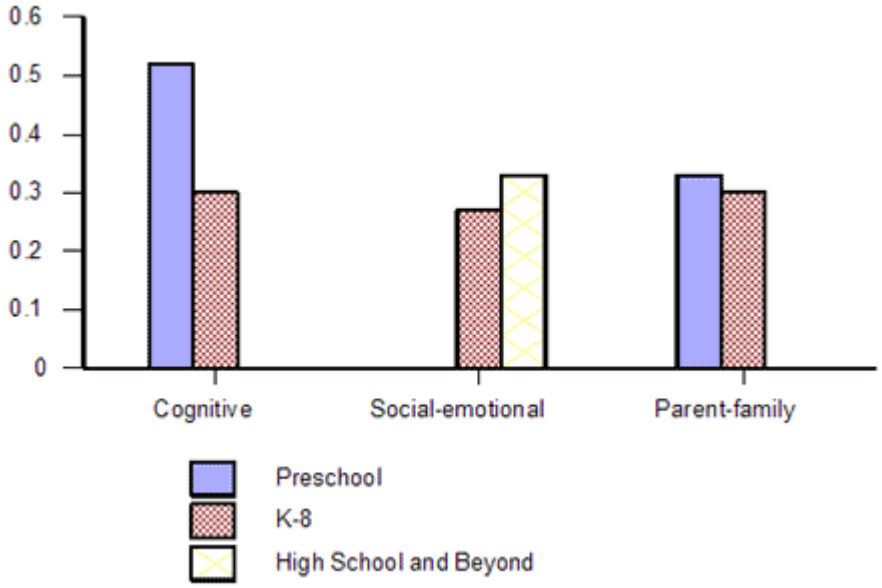

Figure 1. Average effect sizes (d) corrected for sample size for cognitive, social–emotional, and parent–family outcomes at the three time periods. There are no cognitive outcomes at the high school and beyond time period. Social–emotional outcomes for children at preschool are not reported because there were only three studies reporting such outcomes. Parent–family outcomes at high school and beyond are not reported because they are based on only one study.

In Table 3, we present stem-and-leaf displays of the effect sizes, adjusted for sample size, for outcome domains at different time periods: cognitive at preschool, cognitive at K–8, child social–emotional at K–8, child social–emotional at high school and beyond, parent–family at preschool, and parent–family at K–8. Two outliers were identified as approximately 2 standard deviations above the means of their distributions. There was one outlier for cognitive effect sizes at preschool (d = 1.25) and one outlier for parent–family effect sizes at preschool (d = 0.85). Both of these outliers were removed from the data sets before model testing was performed.

Back to article

Table 3
Stem-and-Leaf Displays of Effect Sizes (d) Corrected for Sample Size for Different Outcome Domains at Different Time Periods

	Cognitive		Socialï¿½emotional		Parentï¿½family	
d	Preschool	Kï¿½8	Preschool	XXKï¿½8XX	Preschool	XXKï¿½8XX
1.2						
1.1	5					
1.0	5					
.9	3					
.8						
.7						
.6	0 1		0		5	8
.5	2	6 9	6 8			4

.4	9	2 3 7			
.3	1 6 6	0 0 4	1 3 7	1 9	2 5 6
.2	5 6 8		1 4 6	1	1 6
.1	7	1 4 4 5	4	6 6 8	5 6 6
.0	3	3 5 6 8 9	2 5 7	1 3 5	8 7
−.0	4	2 3 6 7	4 5 6		0 3 5 5
−.01	1	0 0 0 0 3 8	1 5 6	4	2 6

The overall Q test for homogeneity was significant at $p < .001$ for cognitive outcomes at preschool, $Q_t = 102.7$, cognitive outcomes at K–8, $Q_t = 73.2$, social–emotional outcomes at K–8, $Q_t = 49.0$, and parent–family outcomes at K–8, $Q_t = 75.1$. These significant Q_t tests indicate that these distributions are heterogeneous and thus that moderator variable analysis should be performed to examine factors that account for the heterogeneity. The other tests for homogeneity (social–emotional at high school and beyond and parent–family at preschool) were not significant, indicating that these distributions are homogeneous and that further model testing was not necessary. As can be seen from the stem-and-leaf displays in Table 3, the distributions of effect sizes for the outcomes that were significantly heterogeneous show more variability than distributions for the outcomes that were not significantly heterogeneous. No Q_t tests for homogeneity were conducted on cognitive outcomes at high school and beyond, preschool social–emotional outcomes, and parent–family outcomes for high school and beyond because of insufficient sample sizes.

Effect Sizes for Cognitive Outcomes at Preschool and K–8 for Programs With and Without a Preschool Education Program

The first hypothesis was that cognitive effect sizes during preschool and K–8 would be significantly larger for those programs with a preschool education program than for those programs that did not have a preschool education component. During the preschool time period, the average weighted d for programs with a preschool education component ($k = 10$, $d = 0.53$, CI = 0.33–0.73) is much larger than that for programs that did not have a preschool education component ($k = 6$, $d = 0.09$, CI = 0.02–0.15) (see Figure 2). Using the random effects model, we found $Q_b = 64.3$, $p < .001$, indicating that the preschool education component accounted for 63% of the variance in cognitive outcomes (calculated by dividing Q_b by Q_t). At K–8, the average weighted d for cognitive outcomes for programs with a preschool education component ($k = 19$, $d = 0.30$) is still larger than that for programs that did not have a preschool education component ($k = 8$, $d = 0.22$), although the magnitude of the difference is considerably less than for the preschool period. Moreover, because the confidence intervals of the two subgroups overlap, this difference is not statistically significant. These analyses provide some support for the first hypothesis.

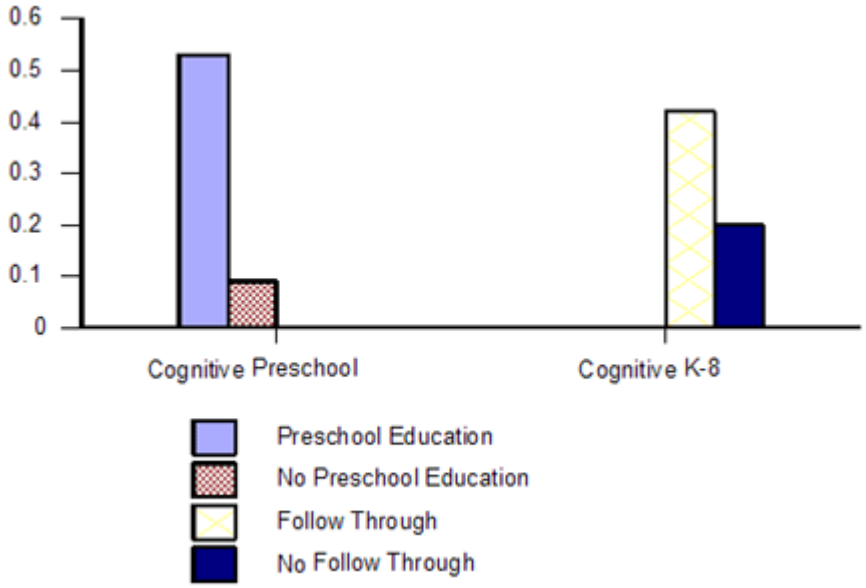

Figure 2. Average weighted effect sizes (*d*) corrected for sample size for cognitive preschool and cognitive K–8 outcomes by presence of educational component for children.

Effect Sizes for Cognitive Outcomes at K–8 for Preschool Programs With and Without Follow-Through Into Elementary School

The second hypothesis was that cognitive effect sizes during the K–8 time period would be larger for programs with a follow-through component than for programs without a follow-through component. The average weighted *d* for programs with a follow-through component ($k = 10$, $d = 0.42$, CI = 0.32–0.51) is significantly larger than that for programs without a follow-through component ($k = 17$, $d = 0.20$, CI = 0.09–0.30) (see Figure 2). Using the random effects model, we found $Q_b = 25.5$, $p < .001$, accounting for 35% of the variance. This analysis provides support for the second hypothesis.

Program Characteristics, Participant Characteristics, and Study Characteristics as Moderators of Cognitive, Social–Emotional, and Parent–Family Outcome Effect Sizes

For the third hypothesis, we reasoned that program characteristics (e.g., length, intensity, comprehensiveness) would be related to effect sizes. We also examined several participant and study characteristics as potential moderators of effectiveness. Two program characteristics, length and intensity of the intervention for children, were found to be related to outcomes. The mean effect size is significantly greater for preschool cognitive outcomes when the length of the intervention was greater than one year ($k = 10$, $d = 0.53$, CI = 0.35–0.72) than when it was less than or equal to one year ($k = 6$, $d = 0.09$, CI = 0.03–0.16). Similar results were found when comparing K–8 social–emotional outcomes for programs greater than one year in length ($k = 15$, $d = 0.27$, CI = 0.16–0.38) with programs less than or equal to one year in length ($k = 4$, $d = 0.06$, CI = 0.01–0.12). Using the random effects model, we found $Q_b = 62.9$, $p < .001$, accounting for 61% of the variance on preschool cognitive outcomes, and $Q_b = 13.4$, $p < .001$, accounting for 27% of the variance on K–8 social–emotional outcomes. See Figure 3 for a graphic depiction of these results. We also examined the scatterplots of length of the intervention for children with these outcomes. For example, the scatterplot in Figure 4 clearly shows that the effect sizes for cognitive outcomes at preschool are lower for programs that were less than 52 weeks in duration than for those that were more than 52 weeks in duration.

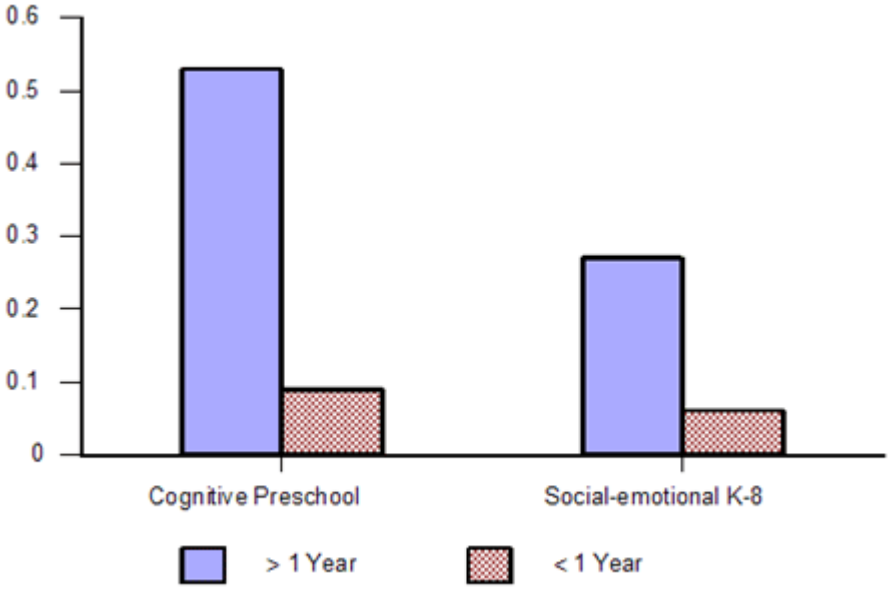

Figure 3. Average weighted effect sizes (*d*) corrected for sample size for different outcomes by length of intervention for child.

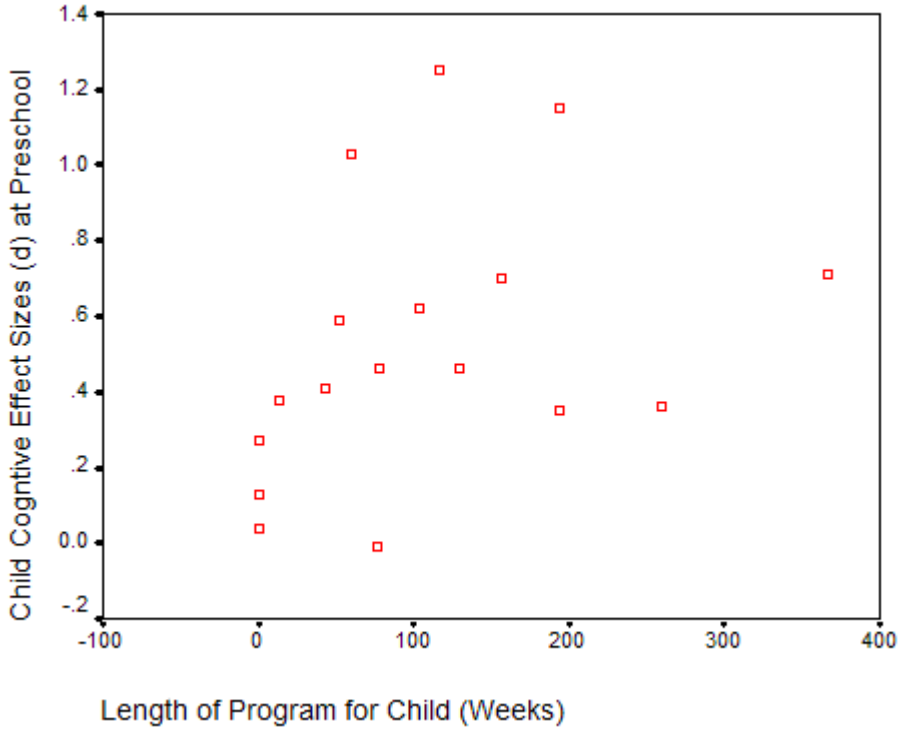

Figure 4. Child cognitive effect sizes (*d*) at preschool by length of program for child.

The mean effect size is also significantly greater for preschool cognitive outcomes when the intensity of the intervention was greater than 300 sessions ($k = 8$, $d = 0.57$, CI = 0.40–0.74) than when it was less than or equal to 300 sessions ($k = 8$, $d = 0.25$, CI = 0.09–0.40). This also holds true for K–8 parent–family outcomes for programs with greater than 300 sessions ($k = 5$, $d = 0.57$, CI = 0.44–0.70) compared with programs that had 300 or fewer sessions ($k = 5$, $d = 0.05$, CI = ?0.03–0.14). Again using the random effects model, we found $Q_b = 64.9$, $p < .001$, accounting for 63% of the variance on preschool cognitive outcomes, and $Q_b = 54.7$, $p < .001$, accounting for 45% of the variance on K–8 parent–family outcomes. No other program characteristics were found to be significant moderators of effect sizes for any of the outcomes. See Figure 5 for a graphic depiction

of these results.

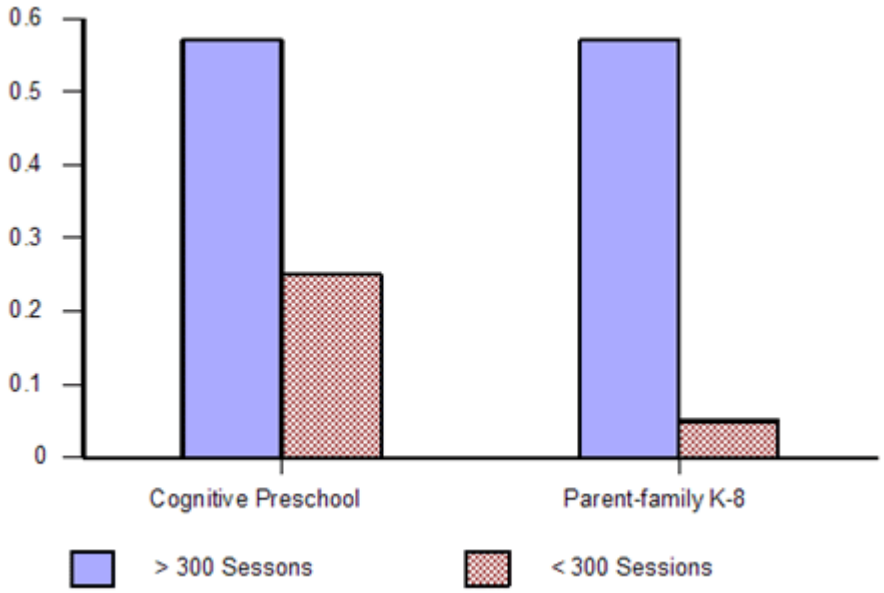

Figure 5. Average weighted effect sizes (d) corrected for sample size for different outcomes by intensity of intervention for child.

We also examined the scatterplots of intensity of the intervention for children with these outcomes. For example, the scatterplot in Figure 6 clearly shows that the effect sizes for parent–family outcomes at K–8 are lower for programs with fewer than 300 sessions than for programs with more than 300 sessions.

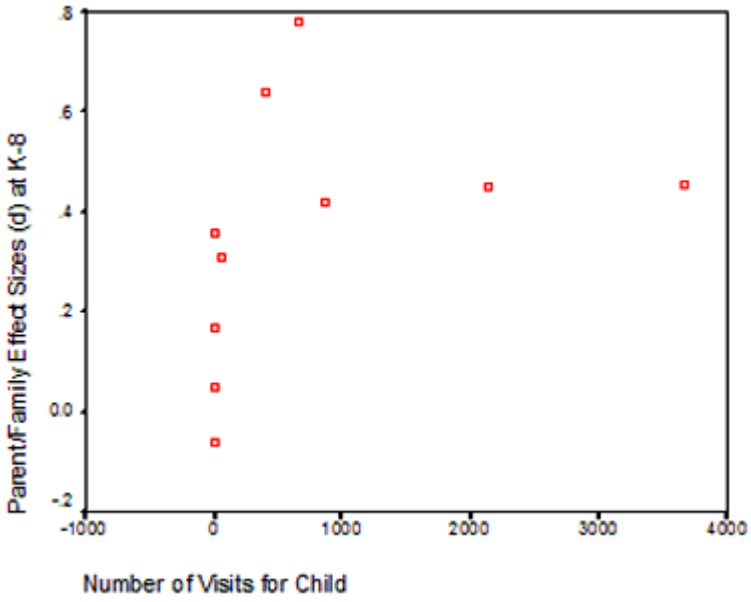

Figure 6. Parent–family effect sizes (d) at K–8 by number of visits for child.

Aside from program characteristics, there was one participant characteristic, ethnoracial background, that was found to be a significant moderator of three of the four outcomes. Interventions that were conducted with predominantly African American children and families show larger effect sizes than interventions for participants with predominantly European American, Mexican American, or other ethnoracial backgrounds. The effect size for preschool cognitive outcomes is significantly larger (Q_b = 32.0, p < .001, 31% of the variance accounted for) for interventions offered to predominantly African American families (k = 9, d = 0.56, CI = 0.41–0.71) than for

those offered to families of other ethnoracial backgrounds ($k = 7$, $d = 0.25$, CI = 0.12–0.39). The effect size for K–8 social–emotional outcomes is also significantly larger ($Q_b = 25.3$, $p < .001$, 52% of variance accounted for) for interventions offered primarily to African American families ($k = 9$, $d = 0.36$, CI = 0.23–0.49) than for those offered to families of other ethnoracial backgrounds ($k = 10$, $d = 0.09$, CI = 0.01–0.16). Finally, the effect size for K–8 parent–family outcomes was significantly greater ($Q_b = 51.8$, $p < .001$, 69% of variance accounted for) for interventions offered predominantly to African American families ($k = 5$, $d = 0.54$, CI = 0.41–0.67) than for those offered to families of other ethnoracial backgrounds ($k = 7$, $d = 0.05$, CI = ?0.03–0.14). These results are graphically depicted in Figure 7.

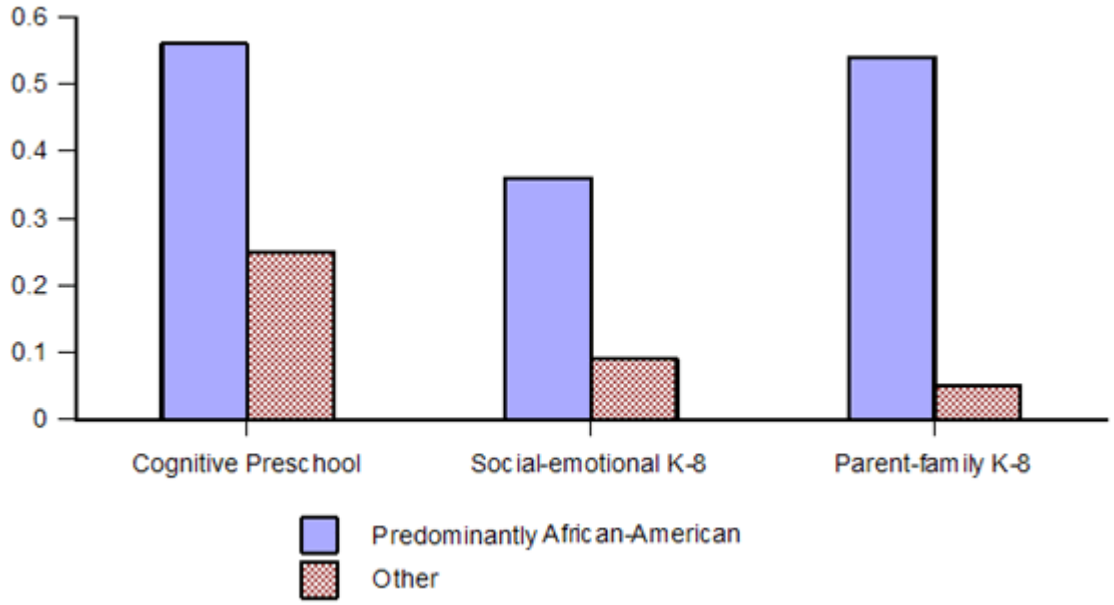

Figure 7. Average weighted effect sizes (*d*) corrected for sample size for different outcomes by ethnoracial background of the study samples.

None of the study characteristics was found to be a significant moderator of any of the outcomes. Neither the total methodology score nor any of the specific methodological components listed in Table 2 (e.g., random assignment, quality of follow-up) were significantly related to any of the outcome domains. Thus, methodological quality is a not a plausible alternative for explaining the impacts of presence of an educational component or length and intensity of the intervention for children on the outcomes.

We also examined whether any of the other significant moderators were related to one another and found one significant relationship. Ethnoracial status was associated with intensity of intervention for the child, $\chi^2(1) = 5.5$, $p < .02$. A total of 74% of studies of predominantly African American participants involved programs with more than 300 sessions compared with 33% of studies of participants from other ethnoracial backgrounds. The point-biserial correlation between ethnoracial status and intensity of the intervention for the child was $r = .35$, $p < .05$.

Fail-Safe n

Meta-analysis is vulnerable to the "file drawer" problem in that studies with nonsignificant findings are unlikely to be published. Failure to include such studies calls into question the findings of a meta-analysis that is based primarily on published studies. The fail-safe *n* determines the number of studies with nonsignificant findings that, if added to the sample, would reduce the combined effect size to a selected level (Orwin, 1983). This was done through the use of the fail-safe *n* formula published by Hedges and Olkin (1985, p. 306). The fail-safe *n* (for $k = 34$ interventions and the overall mean *d* of 0.37) indicates that we would need an additional 29 studies with nonsignificant findings in order to reduce the mean *d* to a small effect size (0.2). It is unlikely that there are 29 missing longitudinal studies of preschool interventions with null results.

Discussion

This meta-analytic review has clearly established that preschool prevention programs have effects on children's cognitive and social–emotional functioning and parent–family wellness that endure through the time that children are in grades K to 8 (i.e., up to nearly 9 years of age). The size of these effects (roughly $d = 0.3$) are in the small to moderate range. When converted to a percentile, this effect size indicates that outcomes for the intervention group sample exceeds 62% of those in the control and comparison group sample (see Lipsey & Wilson, 2001, p. 153, Table 8.1). Moreover, there is further evidence from a smaller number of studies that the impacts of preschool prevention programs on children's social–emotional functioning ($d = 0.33$) continue through the high school years and beyond until the children are, on average, about 18 years of age. This conclusion is consistent with that of a narrative review of the long-term impacts of preschool intervention programs on children (Karoly et al., 1998). Overall, the effect sizes of these long-term impacts are somewhat lower than those reported at postintervention for prevention programs for children and families (Durlak & Wells, 1997; MacLeod & Nelson, 2000). However, given the amount of time that has passed between the preschool period when the programs began and the follow-up to ages 9 and 18, these medium-term and long-term impacts are quite impressive.

The findings of this meta-analysis also support some other conclusions of narrative reviews and go beyond those reviews by clarifying some of the important moderators of program effectiveness for different outcome domains. Like previous reviews, we found that the immediate impacts on children's cognitive development during the preschool period were larger than the medium-term impacts in grades K–8 (Consortium for Longitudinal Studies, 1983; Hubbell et al., 1985). It is not surprising that the most pronounced difference in cognitive outcomes is found during the preschool period, for this is when the life opportunities of disadvantaged children in the intervention and control groups are most different. One group of children has the benefit of a program, while the other does not. When these two groups of children enter school, both enjoy the benefits of education. However, it is important to note that there are enduring cognitive, social–emotional, and parent–family impacts of preschool programs, all of which are similar in magnitude. The "head start" that children receive does make a difference in the long run.

We found support for the hypothesis that programs with a direct teaching component, with a preschool or center-based educational component, have stronger impacts on children's cognitive skills than parent-centered programs without such a component. This finding is consistent with that reported in previous narrative reviews of the literature (Cohen & Radford, 1999; Ramey & Landesman Ramey, 1998). However, our analysis clarifies that this difference is most pronounced during the preschool period. Although cognitive impacts at K–8 are not as large for programs that do not have a preschool education component as for programs that do, they are nevertheless present. This conclusion stands in contrast to that of Cohen and Radford (1999), who asserted that there is little evidence that programs directed at mothers or parents have an impact on children.

We also found support for the hypothesis that programs with a follow-through educational component in elementary school have stronger cognitive impacts at K–8 than programs that do not. This finding is consistent with the conclusion of previous narrative reviews (McLoyd, 1998; Ramey & Landesman Ramey, 1998) and supports the premise of follow-through programs—that ongoing educational intervention can build on preschool education. Although it is known that ongoing educational assistance in the form of special programs and lower teacher–pupil ratios is beneficial, it is not known how much more educational intervention is required before a plateau in cognitive impacts is reached.

Previous reviews have suggested that program length, intensity, breadth, and timing (the age of children when they begin the program) are important moderators of effectiveness (Cohen & Radford, 1999; McLoyd, 1998; Ramey & Landesman Ramey, 1998). We found that four of the distributions of effect sizes were heterogeneous, suggesting that moderator variables might account for this heterogeneity. The two distributions that were homogeneous, social–emotional outcomes at high school and beyond ($k = 10$) and parent–family outcomes at preschool ($k = 7$), had a small number of studies relative to the other distributions. Perhaps the small number of studies accounts for the homogeneity of these distributions.

We examined separately the length and intensity of programs for children and parents. Our findings clearly demonstrate that it is the length and intensity of programs for children that are important moderators of some of

the outcome indicators that were found to be heterogeneous. Moreover, the average weighted effect sizes for those programs that were longer (d = 0.37) and more intense (d = 0.44) are substantially higher than the effect sizes for programs that were shorter (d = 0.11) and less intense (d = 0.18). Longer and more intense programs yielded close to a medium effect size, whereas shorter and less intense programs yielded small effect sizes. Although others (e.g., Ramey & Landesman Ramey, 1998) have suggested that length and intensity of programs for children would be related to children's outcomes, we found that they are related to parent–family outcomes as well. One possible explanation for this finding is that positive changes in the child's behavior have a positive impact on the parents' behavior and the family environment.

Regarding the length and intensity of the program for parents, we did not find any significant relationships between these moderators and outcomes. One possible explanation for this finding is that for the current pool of longitudinal preschool intervention studies, the intensity of the intervention for parents was much less than that for children (the length of the intervention for parents and children was much more similar, with short interventions being less than a year for both parents and children). Thus, it is possible that intensity of the intervention for parents is also very important for positive outcomes for children and parents. Further research on programs that provide more intensive parent–family support interventions is necessary to more adequately determine the role of intensity of intervention for the parents on outcomes.

It has been suggested that programs that begin working with children at a young age have better outcomes than those that begin with children when they are older (Cohen & Radford, 1999; Ramey & Landesman Ramey, 1998). We did not find any evidence to support this assertion. Perhaps as more longitudinal evaluations of preschool prevention programs are completed, such a potential difference may become apparent. Reviewers have also argued that program comprehensiveness or breadth is a moderator of effectiveness, with multicomponent programs being more effective than programs with only one or two components. In our analysis, we operationalized program breadth through the number of program components. We did not find that programs with three or more components were more effective than programs with one or two components on the outcome constructs that we examined. This does not mean that multicomponent programs are not desirable or useful. Such programs may help families meet a variety of different needs, including needs for basic resources such as income and access to other services (Febbraro, 1994). Multicomponent programs may create a web of support for families, rather than having direct impacts on child outcomes, whereas focused programs for children may be necessary to produce outcomes for children. Our measure of program comprehensiveness may have also been inadequate to assess this construct.

The one participant characteristic that had a clear impact on outcomes is ethnoracial status. Across all of the outcomes, effect sizes for studies conducted with predominantly African American children and families (average d = 0.45) are considerably larger than those for studies of children of other ethnoracial backgrounds (average d = 0.14). Why did African American children and families benefit more from preschool intervention programs than children and families from other backgrounds? Many of the programs for African American children and families were conducted in very poor, urban communities in the United States. Perhaps preschool programs benefit those who are most disadvantaged, as McLoyd (1998) has suggested. African American families may be a particularly disadvantaged or oppressed group because of low social class and racism. Also, ethnoracial status was related to intensity of intervention for the child, with African American children disproportionately exposed to the most intense interventions. Because ethnoracial status is related to program intensity, it is difficult to tell at this time which of these factors is most important in accounting for the success of preschool intervention programs.

What are the implications of the findings of this review for programs, policy, and research? First of all, the findings reinforce the assertion that preschool prevention programs are a good example of evidence-based policy. Not only do the results of this analysis clearly demonstrate the beneficial impacts of preschool prevention programs, but cost–benefit and cost-effectiveness evaluations have shown that some of these programs also result in cost savings in the long term (Karoly et al., 1998; Olds, Henderson, Phelps, Kitzman, & Hanks, 1993; Schweinhart & Weikart, 1989). Thus, policies that support preschool prevention programs are a solid social investment and need to be implemented on a more widespread basis.

A second implication concerns the need for an educational component to improve children's cognitive

outcomes. Currently, there has been considerable attention to the development of home visitation programs (Leventhal, 2001), based on the success of Olds' Prenatal/Early Infancy Project in reducing rates of child maltreatment and promoting positive social–emotional outcomes for children and positive social outcomes for mothers. On the basis of this review and the results of an earlier meta-analysis (MacLeod & Nelson, 2000), we know that home visitation programs do have positive impacts on children's cognitive and social–emotional functioning and parent–family outcomes. However, the results of this meta-analysis also demonstrate that educational programs that focus directly on children have more impact on children's cognitive functioning than programs that do not have such a component. Thus, policy-makers, program planners, and practitioners who wish to maximize the impacts of programs on children's cognitive outcomes need to ensure that educational programs focusing directly on children are in place. The findings of this review support the current trend toward multicomponent programs that focus on both parents and children.

A third implication related to that of the need to focus programs on children concerns the strongly interrelated dimensions of length and intensity of the intervention for children. The analyses clearly show that programs less than one year in length and with fewer than 300 sessions had minimal impacts on children. Too often policies are set in place with minimal standards. Programs with minimal standards will produce minimal results. Policies are needed to support programs for children that are longer than one year and that are offered on an intensive basis.

There are two policy implications to the findings regarding ethnoracial status and outcomes. One is that failure to provide preschool prevention programs for low-income African American children and families, for whom they are clearly helpful, is an extremely neglectful social policy. Many years ago, Jensen (1969) argued that preschool education programs for disadvantaged children did not work and should be discontinued. The findings of our review suggest the opposite. Secondly, as McLoyd (1998) argued, there needs to be more research on disadvantaged children from other ethnoracial backgrounds to find ways to make programs more beneficial for children and families from diverse backgrounds. More specifically, there is a need to evaluate the effectiveness of lengthy and high intensity interventions for children from different ethnoracial backgrounds.

Preschool intervention programs have been set in place not only to improve outcomes for children and families but also to deal with the adverse consequences of poverty (McLoyd, 1998). By themselves, preschool prevention programs cannot reduce poverty. There are a variety of social policies dealing with child care, housing, and income maintenance that have been successfully implemented in many Western European countries to reduce child and family poverty (Peters, Peters, Laurendeau, Chamberland, & Peirson, 2001). There has not been sufficient political will in most of North America to implement such powerful family policies. As McLoyd (1998) has observed

> Rather than effect structural changes to remedy poverty and its social ills, U.S. policymakers have overrelied on a variety of social services and programs that call for ameliorating poverty and its attendant problems largely by changing individuals, not structures. (p. 189)

More macro-level policies to reduce and eliminate poverty, provide economic development and employment, and secure affordable housing for low-income citizens are needed to complement the more micro-level programs that we have reviewed (Febbraro, 1994).

With regard to future research, we have already noted that there is a need to examine further programs aimed at children and families from a variety of different ethnoracial backgrounds. Another observation is that very few studies have examined community-level outcomes. Some of the programs that we reviewed had a community development focus and may have had positive impacts on the community. Future research would do well to examine community impacts, as well as impacts on families, parents, and children. One other important area for future research is that of the length and intensity of interventions for parents and families. We found that in general programs for children were much more intensive than those for parents. Further research is needed to examine more adequately the impacts of length and intensity of interventions for parents and families.

There are several limitations of this meta-analytic review. First, there is not a large number of studies of the long-term impacts (high school and beyond) of preschool programs at this point in time. Further longitudinal

investigations of different types of programs in different contexts are needed. A second limitation concerns the grouping of outcome variables. Because of the small number of studies available for the analysis, it was necessary to use broad groupings of constructs (child cognitive, child social–emotional, and parent–family) for the outcome measures reported in the studies. This leads to the "apples and oranges" problem that is endemic to meta-analytic reviews. The broad groupings may gloss over important differences among the outcome measures used. Again, as more longitudinal evaluations of preschool prevention programs become available, researchers will be able to employ more fine-grained outcome constructs.

A third limitation is that we were only able to code variables for which there was sufficient description in the research reports that we reviewed. Some potentially interesting theoretical constructs such as a strengths or empowerment orientation or the cultural sensitivity of the intervention could not be coded because not enough information was provided in the studies about these constructs. A fourth limitation concerns the nature of the control and comparison groups used in the studies that we reviewed. Although we were able to quantify the length and intensity of the intervention in the intervention groups, it was also apparent from reading the articles that children in the control and comparison groups often received some type of intervention as well. However, as with some of other interesting and potentially important variables noted above, there was not enough information to enable us to code the amount of intervention for children in the control and comparison groups.

One final limitation of our meta-analysis is that we treated some continuous variables as categorical variables in our model testing of potential moderators of program success. It would have been preferable to use regression analyses for these continuous variables (e.g., length and intensity of the intervention). Some meta-analysts (e.g., Lipsey & Wilson, 2001) have recently shown how standard software packages, such as SPSS, can be adapted to perform such analyses. All of these limitations need to be addressed in the future both by researchers who are reporting original research on preschool prevention programs and by meta-analysts of such research. Hopefully, in another 10 years there will be a larger number of longitudinal studies whose impacts can be meta-analyzed.

References

Andrews, S. R., Blumenthal, J. B., Johnson, D. L., Kahn, A. J., Ferguson, C. J., Lasater, T. M., Malone, P. E., & Wallace, D. (1982). The skills of mothering: A study of parent child development centers. In F. Horowitz (Ed.), *Monographs of the Society for Research in Child Development*, *47*(6).

Baker, A. J. L., Piotrowski, C. S., & Brooks-Gunn, J. (1999). The Home Instruction Programs for Preschool Youngsters (HIPPY). *The Future of Children*, *9*, 116–133.

Becker, W. C., & Gersten, R. (1982). A follow-up for Follow Through: The later effects of the direct instruction model on children in 5th and 6th grades. *American Educational Research Journal*, *19*, 75–92.

Beller, E. K. (1983). The Philadelphia study: The impact of preschool on intellectual and socioemotional development. In The Consortium for Longitudinal Studies, *As the twig is bent: Lasting effects of preschool programs* (pp. 333–376). Hillsdale, NJ: Erlbaum.

Brooks-Gunn, J. (2003). Do you believe in magic?: What we can expect from early childhood intervention programs. *Social Policy Report: Giving Child and Youth Development Knowledge Away*, *17*, 3–14.

Caruso, G-A. L. (1989). Optimum Growth Project: Support for families with young children. *Prevention in Human Services*, *6*, 123–139.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*, 37–46.

Cohen, N., & Radford, J. (1999). *The impact of early childhood intervention on later life*. Toronto, Ontario, Canada: Hincks-Dellcrest Institute.

Consortium for Longitudinal Studies. (1983). *As the twig is bent: Lasting effects of preschool programs*. Hillsdale, NJ: Erlbaum.

Cullen, K. J. (1976). A six-year controlled trial of prevention of children's behavior disorders. *Journal of Pediatrics*, *88*, 662–666.

Deutsch, M., Deutsch, C. P., Jordan, T. J., & Grallo, R. (1983). The IDS program: An experiment in early and sustained enrichment. In the Consortium for Longitudinal Studies, *As the twig is bent: Lasting effects of preschool programs* (pp. 377–410). Hillsdale, NJ: Erlbaum.

Durlak, J. A. (1995). Understanding meta-analysis. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 319–352). Washington, DC: American Psychological Association.

Durlak, J. A., & Lipsey, M. W. (1991). A practitioner's guide to meta-analysis. *American Journal of Community Psychology*, *19*, 291–332.

Durlak, J. A., & Wells, A. M. (1997). Primary prevention mental health programs for children and adolescents: A meta-analytic review. *American Journal of Community Psychology, 25*, 115–152.

Durlak, J. A., & Wells, A. M. (1998). Evaluation of indicated preventive intervention (secondary prevention) mental health programs for children and adolescents. *American Journal of Community Psychology*, *26*, 775–802.

Evans, E. D. (1985). Longitudinal follow-up assessment of differential preschool experience for low-income minority group children. *Journal of Educational Research*, *78*, 197–202.

Febbraro, A. (1994). Single mothers "at risk" for child maltreatment: An appraisal of person-centred interventions and a call for emancipatory action. *Canadian Journal of Community Mental Health*, *13*(2), 47–60.

Gray, S. W., & Klaus, R. A. (1970). The Early Training Project: A seventh-year report. *Child Development*, *41*, 909–924.

Hedges, L., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Toronto, Ontario, Canada: Academic Press.

Hertzman, C., & Wiens, M. (1996). Child development and long-term outcomes. *Social Science and Medicine*, *43*, 1083–1095.

Honig, A. S., Lally, J. R., & Mathieson, D. (1982). Personal–social adjustment of school children after five years in a family enrichment program. *Child Care Quarterly*, *11*, 138–146.

Hubbell, V. R., McKey, R., Condelli, L., Ganson, H., Barrett, B. J., McConkey, C., & Plantz, M. C. (1985). *The impact of Head Start on children, families and communities: Head Start Synthesis Project* (DHHS Publication No. 85–31193). Washington, DC: Government Printing Office.

The Infant Health and Development Program (1990). Enhancing the outcomes of low-birth-weight, premature infants. *Journal of American Medical Association*, *263*, 3035–3042.

Jensen, A. R. (1969). How much can we boost I.Q. and scholastic achievement? *Harvard Educational Review*, *39*, 1–123.

Karnes, M. B., Schwedel, A. M., & Williams, M. B. (1983). A comparison of five approaches for educating young children from low-income homes. In The Consortium of Longitudinal Studies, *As the twig is bent: Lasting effects*

of preschool programs (pp. 133–169). Hillsdale, NJ: Erlbaum.

Karoly, L. A., Greenwood, P. W., Everingham, S. S., Houbé, J., Kilburn, M. R., Rydell, C. P., Sanders, M., & Chiesa, J. (1998). *Investing in our children: What we know and don't know about the costs and benefits of early childhood interventions*. Santa Monica, CA: RAND.

Keating, D. P., & Hertzman, C. (Eds.). (1999). *Developmental health and the wealth of nations: Social, biological, and educational dynamics*. New York: Guilford Press.

Lambie, D. Z., Bond, J. T., & Weikart, D. P. (1974). *Home teaching with mothers and infants—The Ypsilanti–Carnegie Infant Education Project: An experiment* (Monographs of the High/Scope Educational Research Foundation, No. 6). Ypsilanti, MI: The High/Scope Press.

Lazar, I., & Darlington, R. (1982). Lasting effects of early education: A report from the Consortium for Longitudinal Studies. *Monographs of the Society for Research in Child Development*, *47* (2–3, Serial No. 195).

Lee, V. E., Brooks-Gunn, J., & Schnur, E. (1988). Does Head Start work? A 1-year follow-up comparison of disadvantaged children attending Head Start, no preschool, and other preschool programs. *Developmental Psychology, 24,* 210–222.

Lee, V. E., Brooks-Gunn, J., Schnur, E., & Fong-Ruey, L. (1990). Are Head Start effects sustained? A longitudinal follow-up comparison of disadvantaged children attending Head Start, no preschool, and other preschool programs. *Child Development, 61,* 495–507.

Lee, V., Brooks-Gunn, J., Schnur, E. & Liaw, F. (1990). Are Head Start effects sustained? A longitudinal follow-up comparison of disadvantaged children attending Head Start, no preschool, and other preschool programs. *Child Development*, *61*, 495–507.

Levenstein, P., Levenstein, S., Shiminski, J. A., & Stolzberg, J. E. (1998). Long-term impact of a verbal interaction program for at-risk toddlers: An exploratory study of high school outcomes in a replication of the Mother–Child Home Program. *Journal of Applied Developmental Psychology*, *19*, 267–285.

Leventhal, J. M. (2001). The prevention of child abuse and neglect: Successfully out of the blocks. *Child Abuse and Neglect*, *25*, 431–439.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

MacLeod, J., & Nelson, G. (2000). Programs for the promotion of family wellness and the prevention of child maltreatment: A meta-analytic review. *Child Abuse and Neglect*, *24*, 1127–1149.

MacMillan, H. L., MacMillan, J. H., Offord, D. R., Griffith, L., & MacMillan, A. (1994). Primary prevention of child physical abuse and neglect: A critical review. Part I. *Journal of Child Psychology and Psychiatry, 35*, 835–856.

McCain, M. N., & Mustard, J. F. (1999). *Early years study: Reversing the real brain drain*. Toronto, Ontario, Canada: Ontario Publications.

McLoyd, V. C. (1998). Children in poverty: Development, public policy, and practice. In W. Damon (Editor-in-Chief), I. E. Sigel, & K. A. Renninger (Volume Editors), *Handbook of child psychology, Vol. 4: Child psychology in practice* (pp. 135–208). New York: Wiley.

Meyer, L. A. (1984). Long-term academic effects of the Direct Instruction Project Follow Through. *The Elementary School Journal*, *84*, 380–394.

Miller, L. B., & Bizzell, R. P. (1983a). Long-term effects of four preschool programs: Sixth, seventh, and eighth grades. *Child Development*, *54*, 727–741.

Miller, L. B., & Bizzell, R. P. (1983b). The Louisville Experiment: A comparison of four programs. In Consortium of Longitudinal Studies, *As the twig is bent: Lasting effects of preschool programs* (pp. 171–199). Hillsdale, NJ: Erlbaum.

Miller, L. B., & Dyer, J. L. (1975). Four preschool programs: Their dimensions and effects. *Monographs of the Society for Research in Child Development*, *40*(Serial No. 162, pp. 94–130).

Mrazek, P. J, & Brown, C. H. (2002). Final report: An evidence-based literature review regarding outcomes in psychosocial prevention and early intervention in young children. In C. C. Russell (Ed.), *The state of knowledge about prevention/early intervention* (pp. 42–166). Toronto, Ontario, Canada: Invest in Kids Foundation. Available on the World Wide Web: *http://www.investinkids.ca*.

Nation, M., Crusto, C., Wandersman, A., Kumpfer, K. L., Seybolt, D., Morrissey-Kane, E., & Davino, K. (2003). What works in prevention: Principles of effective prevention programs. *American Psychologist*, *58*, 449–456.

Nelson, G., Laurendeau, M.-C., Chamberland, C., & Peirson, L. (2001). A review and analysis of programs to promote family wellness and prevent the maltreatment of pre-school and elementary school-aged children. In I. Prilleltensky, G. Nelson, & L. Peirson (Eds.), *Promoting family wellness and preventing child maltreatment: Fundamentals for thinking and action* (pp. 220–272). Toronto, Ontario, Canada: University of Toronto Press.

Nelson, G., Prilleltensky, I., & Peters, R. DeV. (2003). Mental health promotion and the prevention of mental health problems in the community. In P. Firestone & W. Marshall (Eds.), *Abnormal psychology: Perspectives* (2nd ed., pp. 462–479). Scarborough, Ontario, Canada: Prentice-Hall.

Olds, D., Eckenrode, J., Henderson, C. R., Kitzman, H., Powers, J., Cole, R., Sidora, K., Morris, P., Pettitt, L. M., & Luckey, D. (1997). Long-term effects of home visitation on maternal life course and child abuse and neglect: Fifteen-year follow-up of a randomized trial. *Journal of the American Medical Association*, *278*, 637–643.

Olds, D., Henderson, C. R., Cole, R., Eckenrode, J., Kitzman, H., Luckey, D., Pettitt, L., Sidora, K., Morris, P., & Powers, J. (1998). Long-term effects of nurse home visitation on children's criminal and antisocial behavior: Fifteen-year follow-up of a randomized controlled trial. *Journal of the American Medical Association*, *280*, 1238–1244.

Olds, D. L., Henderson, C. R., Phelps, C., Kitzman, H., & Hanks, C. (1993). Effects of prenatal and infancy nurse home visitation on government spending. *Medical Care*, *31*, 155–174.

Olds, D. L., & Kitzman, H. (1993). Review of research for pregnant women and parents of young children. *The Future of Children*, *3*, 53–92.

Olds, D. L., & Korfmacher, J. (Eds.). (1998). Prenatal and early childhood home visitation: II. Findings and future directions [Special issue]. *Journal of Community Psychology*, *26*(1).

Orwin, R. G. (1983). A fail-safe *N* for effect size in meta-analysis. *Journal of Educational Statistics*, *8*, 157–159.

Peters, R. DeV., Arnold, R., Petrunka, K., Angus, D. E., Brophy, K., Burke, S. O., Cameron, G., Evers, S., Herry, Y., Levesque, D., Pancer, S. M., Roberts-Fiati, G., Towson, S., & Warren, W. K. (2000). *Developing capacity and competence in the Better Beginnings, Better Futures communities: Short-term findings report*. Kingston, Ontario, Canada: Better Beginnings, Better Futures Research Coordination Unit.

Peters, R. DeV., Peters, J., Laurendeau, M.-C., Chamberland, C., & Peirson, L. (2001). Social policies for

promoting the well-being of Canadian children and families. In I. Prilleltensky, G. Nelson, & L. Peirson (Eds.), *Promoting family wellness and preventing child maltreatment: Fundamentals for thinking and action* (pp. 177–219). Toronto, Ontario, Canada: University of Toronto Press.

Pierson, D. E., Bronson, M. B., Dromey, E., Swartz, J. P., Tivnan, T., & Walker, D. K. (1983). The impact of early education measured by classroom observation and teaching rating of children in kingdergarten. *Evaluation Review*, *7*, 191–216.

Prilleltensky, I., Nelson, G., & Peirson, L. (Eds.). (2001). *Promoting family wellness and preventing child maltreatment: Fundamentals for thinking and action*. Toronto, Ontario, Canada: University of Toronto Press.

Ramey, C. T., Bryant, D. M., Campbell, F. A., Sparling, J. J., & Wasik, B. H. (1988). Early intervention for high risk children: The Carolina Early Intervention Program. In R. H. Price, E. L. Cowen, R. P. Lorion, & J. Ramos-McKay (Eds.), *Fourteen ounces of prevention: A casebook for practitioners* (pp. 32–43). Washington, DC: American Psychological Association.

Ramey, C. T., & Landesman Ramey, S. (1998). Early intervention and early experience. *American Psychologist*, *53*, 109–120.

Rauh, V. A., Achenbach, T. M., Nurcombe, B., Howell, C. T., & Teti, D. M. (1988). Minimizing adverse effects of low birthweight: Four-year results of an early intervention program. *Child Development*, *59*, 544–553.

Reynolds, A. J. (1994). Effects of a preschool plus follow-on intervention for children at risk. *Developmental Psychology*, *30*, 787–804.

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin, 118*, 183–192.

Schwarzer, R. (1996). *Meta-analysis programs*. Unpublished manuscript, Göttingen: Horgrefe. Available on the World Wide Web: http://www.fu-berlin.de/gesund/gesu_engl/meta_e.htm

Schweinhart, L. J., & Weikart, D. P. (1997). The High/Scope Preschool curriculum comparison study through age 23. *Early Childhood Research Quarterly*, *12*, 117–143.

Schweinhart, L. J., & Weikart, D. P. (1989). The High/Scope Perry Preschool study: Implications for early childhood care and education. *Prevention in Human Services, 7,* 109–132.

Schweinhart, L. J., & Weikart, D. P. (1993). Success by empowerment: The High/Scope Perry Preschool Study through age 27. *Young Children*, *49*, 54–58.

Seitz, V., Rosenbaum, L. K., & Apfel, N. H. (1985). Effects of family support intervention: A 10-year follow-up. *Child Development*, *56*, 376–391.

Sprigle, J. E., & Schaefer, L. (1985). Longitudinal evaluation of the effects of two compensatory preschool programs on fourth through sixth grade students. *Developmental Psychology*, *12*, 702–708.

St. Pierre, R. G., Layzer, J. I., Goodson, B. D., & Bernstein, L. S. (1997). *National impact evaluation of the Comprehensive Child Development Program—Final report*. Unpublished report to Administration on Children, Youth and Families, U.S. Department of Health and Human Resources.

Tzuriel, D., Kaniel, S., & Kanner, E. (1999). Effects of the "Bright Start" program in kindergarten on transfer and academic achievement. *Early Childhood Reearch Quarterly*, *14*, 111–141.

Wasik, B. H., Ramey, C. T., Bryant, D. M., & Sparling, J. J. (1990). A longitudinal study of two early intervention

strategies: Project CARE. *Child Development*, *61*, 1682–1696.

Webster-Stratton, C. (1998). Preventing conduct problems in Head Start children: Strengthening parenting competencies. *Journal of Consulting and Clinical Psychology*, *66*, 715–730.

Weissberg, R. P., & Greenberg, M. T. (1998). School and community competence–enhancement and prevention programs. In W. Damon (Editor-in-Chief), I. E. Sigel, & K. A. Renninger (Volume Editors), *Handbook of child psychology: Vol. 4. Child psychology in practice* (pp. 877–954). New York: Wiley.

Wolf, F. M. (1986). *Meta-analysis: Quantitative methods for research synthesis*. Newbury Park, CA: Sage.

Yoshikawa, H. (1994). Prevention as cumulative protection: Effects of early family support and education on chronic delinquency and its risks. *Psychological Bulletin*, *115*, 28–54.

Zigler, E., Taussig, C., & Black, K. (1992). Early childhood intervention: A promising preventative for juvenile delinquency. *American Psychologist*, *47*, 997–1006.

Appendix A
Citations Used in the Coding of Effect Sizes and Study Qualities

Abelson, W. D., Zigler, E., & DeBlasi, C. L. (1974). Effects of a four-year follow-through program on economically disadvantaged children. *Journal of Educational Psychology*, *66*, 756–771.

Achenbach, T. M., Phares, V., Howell, C. T., Rauh, V. A., & Nurcombe, B. (1990). Seven-year outcome of the Vermont Intervention Program for low-birthweight infants. *Child Development*, *61*, 1672–1681.

Andrews, S. R., Blumenthal, J. B., Johnson, D. L., Kahn, A. J., Ferguson, C. J., Lasater, T. M., Malone, P. E., & Wallace, D. (1982). The skills of mothering: A study of parent child development centers. In F. Horowitz (Ed.), *Monographs of the Society for Research in Child Development*, *47*(6).

Baker, A. J. L., Piotrowski, C. S., & Brooks-Gunn, J. (1999). The Home Instruction Programs for Preschool Youngsters (HIPPY). *The Future of Children*, *9*, 116–133.

Becker, W. C., & Gersten, R. (1982). A follow-up for Follow Through: The later effects of the direct instruction model on children in 5th and 6th grades. *American Educational Research Journal*, *19*, 75–92.

Beller, E. K. (1983). The Philadelphia study: The impact of preschool on intellectual and socioemotional development. In The Consortium for Longitudinal Studies, *As the twig is bent: Lasting effects of preschool programs* (pp. 333–376). Hillsdale, NJ: Erlbaum.

Bronson, M. B., Pierson, D. E., & Tivnan, T. (1984). The effects of early education on children's competence in elementary school. *Evaluation Review*, *8*, 615–629.

Campbell, F. A., Goldstein, S., Schaefer, E. S., & Ramey, C. T. (1991). Parental beliefs and values related to family risk, educational intervention, and child academic competence. *Early Childhood Research Quarterly*, *6*, 167–182.

Campbell, F. A., & Ramey, C. T. (1994). Effects of early intervention on intellectual and academic achievement: A follow-up study of children from low-income families. *Child Development*, *65*, 684–698.

Campbell, F. A., & Ramey, C. T. (1995). Cognitive and school outcomes for high-risk African American students at middle adolescence: Positive effects of early intervention. *American Educational Research Journal*, *32*, 743–772.

Caruso, G-A. L. (1989). Optimum Growth Project: Support for families with young children. *Prevention in Human Services*, *6*, 123–139.

Cullen, K. J. (1976). A six-year controlled trial of prevention of children's behavior disorders. *Journal of Pediatrics*, *88*, 662–666.

Cullen, K. J., & Cullen, A. M. (1996). Long-term follow-up of the Busselton six-year controlled trial of prevention of children's behavior disorders. *Journal of Pediatrics*, *129*, 136–139.

Darlington, R. B. (1991). The long-term effects of model preschool programs. In L. Okagaki & R. J. Sternberg (Eds.), *Directors of development: Influences on the development of children's thinking* (pp. 203–215). Hillsdale, NJ: Erlbaum.

Deutsch, M., Deutsch, C. P., Jordan, T. J., & Grallo, R. (1983). The IDS program: An experiment in early and sustained enrichment. In the Consortium for Longitudinal Studies, *As the twig is bent: Lasting effects of preschool programs* (pp. 377–410). Hillsdale, NJ: Erlbaum.

Epstein, A. S., & Weikart, D. P. (1979). *The Ypsilanti–Carnegie Infant Education Project: Longitudinal follow-up* (Monographs of the High/Scope Educational Research Foundation, No. 6). Ypsilanti, MI: The High/Scope Press.

Evans, E. D. (1985). Longitudinal follow-up assessment of differential preschool experience for low-income minority group children. *Journal of Educational Research*, *78*, 197–202.

Gray, S. W., & Klaus, R. A. (1970). The Early Training Project: A seventh-year report. *Child Development*, *41*, 909–924.

Gray, S. W., Ramsay, B. K., & Klaus, R. A. (1983). The Early Training Project: 1962–1980. In the Consortium of Longitudinal Studies, *As the twig is bent: Lasting effects of preschool programs* (pp. 33–69). Hillsdale, NJ: Erlbaum.

Honig, A. S., Lally, J. R., & Mathieson, D. (1982). Personal–social adjustment of school children after five years in a family enrichment program. *Child Care Quarterly*, *11*, 138–146.

The Infant Health and Development Program (1990). Enhancing the outcomes of low-birth-weight, premature infants. *Journal of American Medical Association*, *263*, 3035–3042.

Johnson, D. L., & Walker, T. (1991). A follow-up evaluation of the Houston Parent–Child Development Center: School performance. *Journal of Early Intervention*, *15*, 226–236.

Karnes, M. B., Schwedel, A. M., & Williams, M. B. (1983). A comparison of five approaches for educating young children from low-income homes. In The Consortium of Longitudinal Studies, *As the twig is bent: Lasting effects of preschool programs* (pp. 133–169). Hillsdale, NJ: Erlbaum.

Karnes, M. B., Teska, J. A., & Hodgins, A. S. (1970). The effects of four classroom interventions on the intellectual and language development of 4-year-old disadvantaged children. *American Journal of Orthopsychiatry*, *40*, 58–76.

Klaus, R. A., & Gray, S. W. (1968). The early training project for disadvantaged children: A report after five years. *Monographs of the Society for Research in Child Development, 33*(4), 1–66.

Lally, J. R., Mangione, P. L. & Honig, A. S. (1982). The Syracuse University family development research program: Long-range impact on an early intervention with low-income children and their families. In D. R. Powell (Ed.), *Parent education as early childhood intervention: Emerging directions in theory research, and practice* (pp.

79–104). Norwood, NJ: Ablex Publishing.

Lambie, D. Z., Bond, J. T., & Weikart, D. P. (1974). *Home teaching with mothers and infants—The Ypsilanti–Carnegie Infant Education Project: An experiment* (Monographs of the High/Scope Educational Research Foundation, No. 6). Ypsilanti, MI: The High/Scope Press.

Lee, V., Brooks-Gunn, J., Schnur, E. & Liaw, F. (1990). Are Head Start effects sustained? A longitudinal follow-up comparison of disadvantaged children attending Head Start, no preschool, and other preschool programs. *Child Development*, *61*, 495–507.

Lee, V. E., Schnur, E., & Brooks-Gunn, J. (1988). Does Head Start work? A 1-year follow-up comparison of disadvantaged children attending Head Start, no preschool, and other preschool programs. *Developmental Psychology*, *24*, 210–222.

Levenstein, P., Levenstein, S., Shiminski, J. A., & Stolzberg, J. E. (1998). Long-term impact of a verbal interaction program for at-risk toddlers: An exploratory study of high school outcomes in a replication of the Mother–Child Home Program. *Journal of Applied Developmental Psychology*, *19*, 267–285.

McCarton, C. M., Brooks-Gunn, J., Wallace, I. F., Bauer, C. R., Bennett, F. C., Bernbaum, J. C., Broyles, R. S., Casey, P. H., McCormick, M. C., Scott, D. T., Tyson, J., Tonascia, J., & Meinert, C. L., (1997). Results at age 8 years of early intervention for low-birth-weight premature infants. *Journal of American Medical Association, 277*, 126–132.

Meyer, L. A. (1984). Long-term academic effects of the Direct Instruction Project Follow Through. *The Elementary School Journal*, *84*, 380–394.

Miller, L. B., & Bizzell, R. P. (1983a). Long-term effects of four preschool programs: Sixth, seventh, and eighth grades. *Child Development*, *54*, 727–741.

Miller, L. B., & Bizzell, R. P. (1983b). The Louisville Experiment: A comparison of four programs. In Consortium of Longitudinal Studies, *As the twig is bent: Lasting effects of preschool programs* (pp. 171–199). Hillsdale, NJ: Erlbaum.

Miller, L. B., & Bizzell, R. P. (1984). Long-term effects of four preschool programs: Ninth- and tenth-grade results. *Child Development*, *55*, 1570–1587.

Miller, L. B., & Dyer, J. L. (1975). Four preschool programs: Their dimensions and effects. *Monographs of the Society for Research in Child Development*, *40*(Serial No. 162, pp. 94–130).

Olds, D., Eckenrode, J., Henderson, C. R., Kitzman, H., Powers, J., Cole, R., Sidora, K., Morris, P., Pettitt, L. M., & Luckey, D. (1997). Long-term effects of home visitation on maternal life course and child abuse and neglect: Fifteen-year follow-up of a randomized trial. *Journal of the American Medical Association*, *278*, 637–643.

Olds, D. L., Henderson, C. R., Chamberlin, R., & Tatelbaum, R. (1986). Preventing child abuse and neglect: A randomized trial of nurse home visitation. *Pediatrics*, *78*, 65–78.

Olds, D., Henderson, C. R., Cole, R., Eckenrode, J., Kitzman, H., Luckey, D., Pettitt, L., Sidora, K., Morris, P., & Powers, J. (1998). Long-term effects of nurse home visitation on children's criminal and antisocial behavior: Fifteen-year follow-up of a randomized controlled trial. *Journal of the American Medical Association*, *280*, 1238–1244.

Olds, D. L., Henderson, C. R., & Kitzman, H. (1994). Does prenatal and infancy nurse home visitation have enduring effects on qualities of parental caregiving and child health at 25 to 50 months of life? *Pediatrics*, *93*,

89–98.

Olds, D. L., Henderson, C. R., Chamberlin, R., & Tatelbaum, R. (1988). Improving the life-course development of socially disadvantaged mothers: A randomized trial of nurse home visitation. *American Journal of Public Health*, *78*, 1436–1445.

Olds, D. L., & Korfmacher, J. (Eds.). (1998). Prenatal and early childhood home visitation: II. Findings and future directions [Special issue]. *Journal of Community Psychology*, *26*(1).

Peters, R. DeV., Arnold, R., Petrunka, K., Angus, D. E., Brophy, K., Burke, S. O., Cameron, G., Evers, S., Herry, Y., Levesque, D., Pancer, S. M., Roberts-Fiati, G., Towson, S., & Warren, W. K. (2000). *Developing capacity and competence in the Better Beginnings, Better Futures communities: Short-term findings report*. Kingston, Ontario, Canada: Better Beginnings, Better Futures Research Coordination Unit.

Pierson, D. E., Bronson, M. B., Dromey, E., Swartz, J. P., Tivnan, T., & Walker, D. K. (1983). The impact of early education measured by classroom observation and teaching rating of children in kingdergarten. *Evaluation Review*, *7*, 191–216.

Ramey, C. T., Bryant, D. M., Campbell, F. A., Sparling, J. J., & Wasik, B. H. (1988). Early intervention for high risk children: The Carolina Early Intervention Program. In R. H. Price, E. L. Cowen, R. P. Lorion, & J. Ramos-McKay (Eds.), *Fourteen ounces of prevention: A casebook for practitioners* (pp. 32–43). Washington, DC: American Psychological Association.

Ramey, C. T., & Campbell, F. A. (1984). Prevention education for high-risk children: Cognitive consequences of the Carolina Abecedarian Project. *American Journal of Mental Deficiency*, *88*, 515–523.

Ramey, C. T., & Campbell, F. A. (1991). Poverty, early childhood education, and academic competence: The Abecedarian experiment. In A. C. Huston (Ed.), *Children in poverty: Child development and public policy* (pp. 190–221). Cambridge, England: Cambridge University Press.

Rauh, V. A., Achenbach, T. M., Nurcombe, B., Howell, C. T., & Teti, D. M. (1988). Minimizing adverse effects of low birthweight: Four-year results of an early intervention program. *Child Development*, *59*, 544–553.

Reynolds, A. J. (1992). Mediated effects of preschool intervention. *Early Education and Development*, *3*, 139–164.

Reynolds, A. J. (1994). Effects of a preschool plus follow-on intervention for children at risk. *Developmental Psychology*, *30*, 787–804.

Reynolds, A. J., Chang, H., & Temple, J. A. (1998). Early childhood intervention and juvenile delinquency: An exploratory analysis of the Chicago Child–Parent Centers. *Evaluation Review*, *22*, 341–372.

Reynolds, A. J., Mehana, M., & Temple, J. A. (1995). Does preschool intervention affect children's perceived competence? *Journal of Applied Developmental Psychology*, *16*, 211–230.

Reynolds, A. J., & Temple, J. A. (1995). Quasi-experimental estimates of the effects of a preschool intervention: Psychometric and econometric comparisons. *Evaluation Review*, *19*, 347–373.

Reynolds, A. J., & Temple, J. A. (1998). Extended early childhood intervention and school achievement: Age thirteen findings from the Chicago Longitudinal Study. *Child Development*, *69*, 231–246.

Royce, J. M., Darlington, R. B., & Murray, H. W. (1983). Pooled analyses: Findings across studies. In Consortium for Longitudinal Studies, *As the twig is bent: Lasting effects of preschool programs* (pp. 411–459).

Hillsdale, NJ: Erlbaum.

Royce, J. M., Lazar, I., & Darlington, R. B. (1983). Minority families, early education, and later life chances. *American Journalof Orthopsychiatry*, *53*, 706–720.

Schweinhart, L. J., & Weikart, D. P. (1983). The effects of the Perry Preschool Program on youths through age 15—A summary. In Consortium of Longitudinal Studies, *As the twig is bent: Lasting effects of preschool programs* (pp. 71–101). Hillsdale, NJ: Erlbaum.

Schweinhart, L. J., Weikart, D. P., & Larner, M. B. (1986). Consequences of three preschool curriculum models through age 15. *Early Childhood Research Quarterly*, *1*, 15–45.

Seitz, V., Apfel, N. H., & Efron, C. (1978). Long-term effects of early intervention: The New Haven project. In B. Brown (Ed.), *Found: Long-term gains from early intervention* (pp. 79–109). Boulder, CO: Westview Press.

Seitz, V., Rosenbaum, L. K., & Apfel, N. H. (1985). Effects of family support intervention: A 10-year follow-up. *Child Development*, *56*, 376–391.

St. Pierre, R. G., Layzer, J. I., Goodson, B. D., & Bernstein, L. S. (1997). *National impact evaluation of the Comprehensive Child Development Program—Final report*. Unpublished report to Administration on Children, Youth and Families, U.S. Department of Health and Human Resources.

St. Pierre, R. G., & Layzer, J. I. (1999). Using home visits for multiple purposes: The Comprehensive Child Development Program. *The Future of Children*, *9*, 134–151.

Sprigle, J. E., & Schaefer, L. (1985). Longitudinal evaluation of the effects of two compensatory preschool programs on fourth through sixth grade students. *Developmental Psychology*, *12*, 702–708.

Tzuriel, D., Kaniel, S., & Kanner, E. (1999). Effects of the "Bright Start" program in kindergarten on transfer and academic achievement. *Early Childhood Reearch Quarterly*, *14*, 111–141.

Wasik, B. H., Ramey, C. T., Bryant, D. M., & Sparling, J. J. (1990). A longitudinal study of two early intervention strategies: Project CARE. *Child Development*, *61*, 1682–1696.

Webster-Stratton, C. (1998). Preventing conduct problems in Head Start children: Strengthening parenting competencies. *Journal of Consulting and Clinical Psychology*, *66*, 715–730.

Weikart, D. P. (1989). Early childhood education and primary prevention. *Prevention in Human Services*, *6*, 285–306.

Weikart, D. P. (1998). Changing early childhood development through educational intervention. *Preventive Medicine*, *27*, 233–237.

Appendix B
Citations Used Only in Coding Study Qualities

Beller, E. K. (1974). Impact of early education on disadvantaged children. In S. Ryan (Ed.), *A report on longitudinal evaluations of preschool programs* (DHEW Publication No. 75–24, pp. 15–48). Washington, DC: Department of Health, Education, and Welfare.

Day, M. C., & Parker, R. K. (1972). *The preschool in action*. Boston, MA: Allyn and Bacon.

Johnson, D. L. (1990). The Houston Parent–Child Development Center Project: Disseminating a viable program for enhancing at-risk families. *Prevention in Human Services*, *7*, 89–108.

Ramey, C. T., Bryant, D. M., Wasik, B. H., Sparling, J. J., Fendt, K. H., & LaVange, L. M. (1992). Infant health and development program for low-birth-weight, premature infants: Program elements, family participation, and child intelligence. *Pediatrics*, *3*, 455–468.

Schweinhart, L. J., & Weikart, D. P. (1989). The High/Scope Perry Preschool study: Implications for early childhood care and education. *Prevention in Human Services*, *7*, 109–132.

Schweinhart, L. J., & Weikart, D. P. (1997). The High/Scope Preschool curriculum comparison study through age 23. *Early Childhood Research Quarterly*, *12*, 117–143.

Footnote

[1] First, for the Illinois, Louisville, and Washington interventions, we summed the effect sizes across the multiple interventions within each of these studies. Second, because two different designs were used to evaluate the three Better Beginnings sites (a within-site baseline-focal cohort design and a longitudinal comparison group), we only used the results from the baseline-focal cohort design because the comparison groups were within the sites and independent of one another. Third, for other studies that compared interventions of differing scope and intensity, we selected those interventions or comparison groups that maximized the differences between the intervention and control group (i.e., the strongest interventions or the purest no-intervention controls). Thus, we included only the 8-year Abecedarian intervention (not the 3-year or 5-year interventions), the family education and child development center intervention for Project CARE (not the family-education-only intervention), the nursery, kindergarten, and Grade 1 intervention for the Philadelphia Project (not the nursery and kindergarten program), the comparison of the Perry Preschool with the no-intervention control group (not the DISTAR alternative intervention comparison group), and the 13 Head Start programs for the Portland and Trenton study (not the "6 other preschool" group). For the Chicago Child Parent Centers (CPC), we used the analyses that were most frequently reported across time periods in the different articles—the CPC preschool versus no-preschool comparison.