Chapter Two

TARGETED EARLY INTERVENTION PROGRAMS AND THEIR BENEFITS

In this chapter, we assemble the evidence regarding the benefits of targeted early childhood intervention for program participants. We do this principally by reviewing the evolution of targeted early intervention programs from the initial efforts in the 1960s to the present, highlighting the findings from formal evaluations of several of the most noteworthy programs. We conclude by summarizing the evidence available across programs.

THE ORIGINS OF EARLY INTERVENTION

Early intervention as we know it today is the consequence of over 300 years of social, political, and academic influences. Many of the fundamental theories upon which today's interventions are founded have their origins in early concepts of childhood development.

The earliest modern influences are the writings of European philosophers of the Enlightenment who were among the first to envision childhood as a separate period in life subject to unique developmental influences (Aries, 1962). The debate concerning "nature versus nurture" can be traced to the theories of Locke and others who considered the child a tabula rasa, or clean slate, upon which the environment and upbringing were entirely influential (Meisels and Shonkoff, 1990). This is in contrast to the belief that children are born with genetically predetermined attributes that the environment can do little to change.

This debate continues to echo today through academia and society at large. Many theorists now consider a hybrid of these views to be the most accurate depiction of early childhood potential: While children are born with certain innate capabilities, they are subject to the influences of environment and parenting, and it is a combination of these factors that contributes to later functioning (Carey and Gelman, 1991).

The social construct of childhood evolved in the 19th century into a family-centered model as the industrial revolution created the possibility of a childhood sheltered from adult responsibilities (Aries, 1962). European influences on early childhood theory were prominent at this time, with the emergence of an organized early education curriculum in Germany designed by Friedrich Froebel. These first "kindergarten" classes were designed to promote religious value–based child development through structured play activities (Meisels and Shonkoff, 1990).

Child development theory was also influenced by the evolution of the social sciences and the application of scientific methods of observation, data collection, and analysis to studies of human populations. The scientific approach drew support from the positivist theorists of this period, who believed that there were universal truths about society that once revealed would lead to the resolution of society's problems (Lubeck, 1995). The turn of the century saw the first theories and tests of intelligence developed by Binet and Simon in Paris, and their use to distinguish intellectual capabilities of children (Flanagan, Genshaft and Harrison, 1997). These tests rapidly gained popularity within American culture.

The importation of the kindergarten curriculum to the United States coincided with dramatic changes in the urban American landscape as new waves of immigrants began arriving in American cities. While the middle classes embraced kindergarten as a means to improve child development, the poor urban environment and the growing heterogeneity of the population led social activists to promote the kindergarten curriculum as an important service for immigrants (Ramey, Dorval, and Baker-Ward, 1983). It was hoped that kindergarten would compensate for economic disadvantage in early childhood experiences, promote cultural assimilation, and provide safe

child care for working mothers in urban industrial slums (Cohen, 1996).

This call for targeting early intervention to an "at-risk" population is one of the first encroachments upon the Anglo-American tenet of *parens patriae.* This principle considers intervention in a child's upbringing to be justifiable only when parents cannot or will not care for their children (Melton and Megan, 1993). Advocates of early education for immigrants felt that urban conditions were sufficiently deficient to require intervention in the care and upbringing of these children. Early interventionists, under the rubric of "child saving," advocated center-based early education partly to remove children from homes considered economically and culturally impoverished (Swadener and Lubeck, 1995; Zigler, 1990).

Soon after the kindergarten movement made its way to the United States, the "nursery school" curriculum was designed in London by Rachel and Margaret McMillan. In its first incarnation, the nursery school was also a targeted program that began as an urban health clinic (Meisels and Shonkoff, 1990). As it evolved, the nursery school became a comprehensive program designed to improve both education and health. In the United States, the nursery school became a desired "universal," but privately funded, service purchased by middle and upper class families as a means of enabling children to achieve future success (Cohen, 1996).

Continued efforts in the late 19th century to improve urban conditions for children included the opening of "settlement houses" in impoverished communities. Lillian Wald, who founded the Henry Street Settlement in New York City in 1893, and Jane Addams, who established Hull House in Chicago in 1889, were prominent social activists whose work led to a call for federal intervention in the plight of poor children (Hutchins, 1994). This work culminated in the White House Conference of 1909 called by President Theodore Roosevelt, which in turn led to the creation of the federal Children's Bureau in 1912, the first federal agency concerned exclusively with the health and well-being of children (Parmelee, 1994).

The research on infant mortality sponsored by the Children's Bureau led to the Sheppard-Towner Act of 1921, which established maternal and child health services in each state funded through federal grants-

in-aid matched by state funds (Hutchins, 1994). This was the first expression of a national child welfare policy establishing the responsibility of states and communities to provide services to improve early childhood. This act was quite controversial and was opposed by the Roman Catholic Church, the U.S. Public Health Service, and the American Medical Association (AMA) (Hutchins, 1994).[1]

Until the 1920s, there was little scientific data on the normal developmental and physical milestones of childhood. The work of Arnold Gesell, a strongly nativist theoretician, established the first parameters of normal and abnormal child development (Gesell and Thompson, 1934). His research provides the foundation for measurement tools used to assess development today. Other movements in child development research included behavioral theorists, such as John B. Watson, who, unlike Gesell, believed that the most important determinants of development were environmental (Watson, 1913). The White House Conference on Child Health and Protection of 1930 was called to draw upon the growing body of research of the 1920s on children's normative growth and development. The Children's Charter—19 statements describing the education, health, welfare, and protection conditions required for optimal child development—was the landmark product of the White House conference.

The 1930 conference paved the way for expansion of maternal child services in the enactment of Title V of the Social Security Act of 1935. This act authorized grants-in-aid to states for expanded child welfare services, and for maternal and child health programs. It recognized the dependence of children on society for their welfare and acknowledged a responsibility for improving social as well as health conditions (Hutchins, 1994).

The federal role in provision of non-health-related child services grew during World War II as large numbers of middle- as well as working-class women entered the workforce. Federally supported child care and nursery schools were created under the Lanham Act during the 1940s to care for children of working women. The emphasis on early education in child care became more prominent

[1]It was the AMA's opposition that led a group of physicians belonging to the new profession of pediatrics to establish the American Academy of Pediatrics (AAP) in 1930 (Hutchins, 1994).

with the increase in middle-class use of these services. After the war, however, federal subsidizes for nursery schools virtually disappeared as many of these women left the workforce. Nursery schools in the 1950s once again became the privilege of the middle classes (Meisels and Shonkoff, 1990).

In the post–World War II era, developmental research was strongly influenced by psychoanalytic theory and the examination of deprivation in early childhood. Ainsworth, Bowlby, Anna Freud, and others developed the first theories of mother-child attachment in the 1950s (Ainsworth et al., 1978; Bowlby, 1969; Burlingham and Freud, (1949a, 1949b, 1950). Erikson advanced a theory of normal emotional development (Erikson, 1985). The study of causes of mental retardation became prominent as Piaget's work on normal cognitive development became well known (Piaget, 1947; Piaget and Inhelder, 1947).

Postwar affluence may have contributed to the increasing focus of the lay public on child rearing and development. The popular conceptualization of genetic predetermination of personality and intelligence began to give way under the influence of the academic writings of Donald Hebb, J. McVicker Hunt, and Benjamin Bloom. Hebb's neuropsychological theory examined the dependence of learning later in life on the quality and quantity of early experiences (Hebb, 1949). Hunt emphasized the importance of the first 3 years of life and argued that early environmental influences were the most powerful determinants of child outcomes (Hunt, 1961). Bloom theorized that 50 percent of intellectual development occurred by age 4, implying that a child's development will be most affected by environmental influences in this period of rapid intellectual development (Bloom, 1964). The theoretical groundwork was laid for the debates of the following three decades regarding the causes of abnormal child development and the optimal means of preventing and improving it.

FROM THEORY TO PRACTICE: DETERMINING WHICH INTERVENTIONS WORK

With a twentyfold increase in federal spending on social programs from 1950 to 1979, policymakers and social scientists alike appreci-

ated the need to study how programs were implemented and what results they produced (Shadish, Cook, and Leviton, 1995). Evaluation research has filled the need for a systematic and scientific approach to assessing the conceptualization, design, implementation, and utility of interventions (Rossi and Freeman, 1993).

The difficulty in drawing conclusions about the success of early intervention programs is that the analytic methodology of evaluation has evolved side by side with the programs targeted for study. Our ability to judge the effectiveness of early programs is frequently constrained by the limitations of early outcome measures. For example, early researchers generally assumed that program effects would be specific and immediate, and evaluations were not designed to measure broad or long-term outcomes (Shadish, Cook, and Leviton, 1995).

There are also intrinsic incompatibilities between social science evaluation methods and social service delivery (Weiss, 1972). While properly implemented randomized control trials remain the ideal in terms of evaluation methodology, evaluators must frequently adjust their research methods to the priorities and time constraints of the program they are attempting to study. Experimental-design considerations, such as specific "treatments" and randomized assignment of participants, often do not readily accommodate the service function of intervention programs. Implementing randomized trials and other evaluative methods also takes money away from service provision, and it may be considered unethical to deny control groups access to services. In addition, even with random assignment in longitudinal studies, differential and possibly nonrandom attrition between treatment (or experimental) and control groups may also bias findings from experimental designs, especially when sample sizes are small.

Many of the criticisms directed at early intervention evaluations have to do with the drawing of causal inferences from quasi-experimental and nonexperimental study designs that do not employ random assignment (Reynolds and Temple, 1995).[2] Nonrandom assignment

[2]Nonexperimental studies are those in which there is no comparison group. A comparison group is a set of children who are similar to the program participants in as many ways as possible except that they do not participate. The best way to ensure a

may not control for differences in observed and unobserved characteristics between children who participated in an intervention and the comparison group. This may lead to potential bias in estimates of the program's effectiveness and the necessity of cautious interpretation of study results. Because there is often no single or simple way to analyze nonexperimental and quasi-experimental data, controversy surrounds the outcomes claimed for early intervention and will continue to do so (General Accounting Office (GAO), 1997). There are, however, statistical techniques for evaluating quasi-experimental and nonexperimental studies that researchers may use to make estimates of program effects. Newer techniques reduce selection bias by incorporating selection into the analytic model. Reanalysis of previously studied programs has found in some cases favorable effects due to early intervention programs where previously none were recognized (see, for example, Currie and Thomas, 1995, and Reynolds and Temple, 1995).

The political context of social service programs is also an important factor in program and evaluation design, completion, and reporting (Weiss, 1973). Social programs and their evaluations are subject to the ideological debate that gives rise to them, and evaluations that do not support a political view may not find a public venue. Basic political assumptions of service providers may also limit the ability to fully assess the limitations or successes of certain programs.

Finally, most programs have multiple stakeholders who require different questions to be answered by evaluation studies, or who may have little if any interest in measured outcomes (Shadish, Cook, and Leviton, 1995). The literature available on many early intervention programs reflects the diversity of research interests and reinforces the difficulty in drawing common conclusions from different intervention programs and their evaluations.

Together, these issues of study and analysis design raise concerns about the predictive value of the research on early intervention (GAO, 1997). Other issues concern social science research in general.

---

matched comparison group is to recruit a set of children for the program and randomly assign some to the program; the remainder constitute the comparison (or control) group. Quasi-experimental designs are those including a comparison group that has been chosen on the basis of matched characteristics but not random assignment.

An important bias of the literature is that mainly favorable outcomes are published, so it may not reflect a wider, possibly negative, body of evidence about the area of concern. However, there may also have been many programs that have contributed favorably to the health and well-being of children and families but were not reported in the literature. Many programs do not have an evaluation component, or many may have lost funding or faced other crises before their evaluations were completed. Finally, long-term follow-up evaluations necessitate a lapse of years between intervention and publication of results. In that time, the world may have changed enough to make it imprudent to uncritically apply the lessons learned to new interventions (Lazar and Darlington, 1982).

HISTORICAL REVIEW OF TARGETED EARLY INTERVENTION PROGRAMS

Many excellent and comprehensive reviews of the early intervention literature have been written since the 1960s, when early intervention programs began to proliferate (for recent reviews, see Meisels and Shonkoff, 1990; Seitz, 1990; Benasich, Brooks-Gunn, and Clewell, 1992; Olds and Kitzman, 1993; Barnett, 1995; Yoshikawa, 1995; Guralnick, 1997; Reynolds et al., 1997). These reviews can be differentiated based on the criteria the authors used for inclusion of studies in their analysis. Many reviews are oriented to a specific type of intervention or a specific age group or type of outcome. Other reviews use methodological criteria—including design, sample size, attrition, and length of follow-up period—to delineate the literature.

In the remainder of this chapter, we draw on the early intervention literature and previous reviews to highlight some of the most prominent targeted early intervention programs, with particular attention to studies potentially relevant to cost-savings analysis. In particular, we focus on ten programs:

- Early Training Project
- High/Scope Perry Preschool Project
- Project Head Start
- Chicago Child-Parent Center (CPC) and Expansion Program

- Houston Parent-Child Development Center (PCDC)
- Syracuse Family Development Research Program (FDRP)
- Carolina Abecedarian
- Project CARE (Carolina Approach to Responsive Education)
- Infant Health and Development Project (IHDP)
- Elmira Prenatal/Early Infancy Project (PEIP).

These programs share a common aim: to improve child health and development by providing socioeconomically disadvantaged children and their families with various services and social supports during part or all of the period of early childhood. Although most of these programs are considered "model programs," two have been implemented on a large scale—Project Head Start and the Chicago Child-Parent Center and Expansion Program—and the Infant Health and Development Project was implemented at eight sites. The remaining programs were implemented at one site, although replications in other settings are under way for at least one program. Taken together, these ten programs have been some of the most influential in establishing an understanding of the effects of targeted early intervention programs.

We focus on these studies because they also meet several other criteria we believe are important for providing more-reliable estimates of both short-run and long-run program effects: experimental design, preferably with randomized assignment to treatment and control groups; a sample size of 50 children or more in treatment plus controls; a follow-up period, preferably past the period of program intervention; and less than 50 percent attrition at follow-up.[3] The ten programs that we review here do not exhaust those that would meet the criteria we list above.[4] Rather, our goal is to illustrate the range of

---

[3]Besides excluding studies with nonexperimental designs, small samples, no follow-up, or high attrition, we also do not focus on analyses of multiple studies. For example, the Consortium for Longitudinal Studies (Lazar and Darlington, 1982) conducted a collaborative follow-up for several early childhood programs, including the Early Training Project and the Perry Preschool program, which we review here.

[4]For example, Barnett (1995) reviews a longer list of programs using similar criteria, although his review does not include Project CARE, the Infant Health and Development Project, or the Prenatal/Early Infancy Project. By excluding other published

interventions that have been studied, highlighting the types of effects that have been measured for these representative programs. In the concluding section of this chapter, we take account of the broader literature and highlight any differences between the inferences we draw based on the ten studies featured here and those based on a more comprehensive review of the literature.

Table 2.1 provides information on the characteristics of the programs we review, including the place and time period of implementation, the characteristics of the intervention (i.e., target group; ages of participation; and the focus, mode, and nature of services provided), and the features of the evaluation (i.e., whether a randomized controlled experiment, initial and follow-up sample sizes, and the ages of the children at follow-up). The table and the discussion of the programs that follows highlight the diversity of these interventions. In part, the diversity arises from the differences in program objectives—program components are designed to mediate such different stressors in early childhood as low income, poor health, or lack of cognitive stimulation or emotional support. Design features that are affected by program objectives include how children are targeted for intervention, which members of the family receive program services or treatments, and the nature and mode of service delivery. As summarized in Tables 2.2 and 2.3, the variation in program objectives also leads to differences in the outcomes measured for children and their parents.

First, targeted early intervention programs may select participants using different criteria, with different definitions of which children and families are most in need—or most at risk of being exposed to stressors that might compromise development. Impediments to optimal child outcomes include impaired emotional relationships; insufficient cognitive stimulation; resource deficiencies associated with low income, poor nutrition, or inadequate access to health care; and, more controversially, linguistic and cultural factors. In the studies we review, the criteria for identifying children for treatment include being in a family with low socioeconomic status (SES), having a low IQ, or being born prematurely with a low birthweight.

---

studies from our review, we do not mean to imply that their results are necessarily biased.

Table 2.1

Features of Selected Targeted Early Intervention Programs

Program Years of Operation Site	Intervention					Evaluation Design			
	Target	Ages of Participants	Focus	Mode	Content	Random Assignment	Initial Sample Size	Sample Size at Final Follow-Up	Ages at Follow-Up
Early Training Project[a] 1962–1965 Murfreesboro TN	Low SES	Entry: 4 to 5 years Exit: 6 years	Child	Center/ home	Summer part-day preschool program. Home visits.	Yes	E=44 C=21	E=36 C=19	6, 7, 8, 10, 16–20
High/Scope Perry Preschool Project[b] 1962–1967 Ypsilanti, MI	Low SES and low IQ scores	Entry: 3 to 4 years Exit: 5 years	Child	Center/ home	School-year part-day preschool program. Home visits.	Yes	E=58 C=65	E=58 C=63	5–11, 14, 15, 19, 27
Project Head Start 1965–present multiple	Low SES	Entry: 3 years Exit: 4 years	Child	Center/ home	Preschool program. Home visits.	NA	NA	NA	NA
Chicago CPC[c] 1967–present Chicago, IL	Low SES	Entry: 3 to 4 years Exit: 6 to 9 years	Child/ parent	Center	Preschool: Half-day school-year program. School-age: Kindergarten and primary (to 3rd grade) programs.	No; statistical controls for nonrandom participation	E=1,150 C=389	E=878 C=286	9, 10, 11, 14
Houston PCDC[d] 1970–1980 Houston, TX	Low SES	Entry: 1 (HV); 2 (center) Exit: 3 years	Child/ parent	Home/ center	Home visits. Part-day child care. Center-based program for parents.	Yes	E=90 C=201	School data: E=50 C=87 Achievement tests: E=39 C=76	3, 4–7, 8–11

Table 2.1—continued

Program Years of Operation Site	Intervention					Evaluation Design			
	Target	Ages of Participants	Focus	Mode	Content	Random Assignment	Initial Sample Size	Sample Size at Final Follow-Up	Ages at Follow-Up
Syracuse FDRP[e] 1969–1975 Syracuse , NY	Low SES	Entry: last trimester (HV); 6 months (center) Exit: 5 years	Child/ parent	Home/ center	Home visits. Part-day (6–15 months) to full-day (15–60 months) year-round family-style day care.	No; matched comparison group selected at 36 months	E=108 C=108	E=65 C=54	5, 6, 15
Carolina Abecedarian[f] 1972–1985 one site in NC	High score on high-risk index	Entry: 6 weeks to 3 months Exit: 5 to 8 years	Child/ parent	Center	Preschool: full-day year-round center-based educational day care. School-age: parent program.	Yes	E=57 C=54	E=48 C=44	8, 12, 15
Project CARE[g] 1978–1984 one site in NC	High score on high-risk index	Entry: 4 weeks (HV); 6 weeks to 3 months (center) Exit: 5 years	Child/ parent	Home/ center	Home visits and full-day year-round center-based educational day care (E1). Home visits only (E2).	Yes	E1=17 E2=25 C=23	E1=14 E2=23 C=22	4-1/2
IHDP[h] 1985–1988 8 sites	Premature and low birthweight	Entry: birth (HV); 1 year (center) Exit: 36 months (adjusted for prematurity)	Child/ parent	Home/ center	Home visits. Full-day year-round center-based educational day care.	Yes	E=377 C=608	E=336 C=538	3, 5, 8

Table 2.1—continued

Program Years of Operation Site	Intervention					Evaluation Design			
	Target	Ages of Participants	Focus	Mode	Content	Random Assignment	Initial Sample Size	Sample Size at Final Follow-Up	Ages at Follow-Up
Elmira PEIP[i] 1978–1982 Elmira, NY	First births to young, single or low-SES mothers	Entry: up to 30th week of gestation Exit: 2 years	Parent	Home	Home visits by trained nurses.	Yes	E=116 C=184	E=97 (mothers), 94 (children) C=148 (mothers), 144 (children)	3, 4, 15

NOTES: Age references are to the age of the focal child. E = experimental; C = control; HV = home visits; NA = not applicable.

[a]Gray and Klaus (1970); Gray and Ramsey (1982); Gray, Ramsey, and Klaus (1982); Lazar and Darlington (1982).

[b]Weikart, Bond, and McNeil (1978); Schweinhart and Weikart (1980); Berrueta-Clement et al. (1984); Schweinhart et al. (1993).

[c]Reynolds (1994, 1997); Reynolds and Temple (1995); Reynolds, Chang, and Temple (1997).

[d]Johnson et al. (1974); Andrews et al. (1982); Johnson and Breckenridge (1982); Johnson and Walker (1991).

[e]Honig and Lally (1982); Lally, Mangione, and Honig (1988).

[f]Ramey, Dorval, and Baker-Ward (1983); Ramey and Campbell (1984, 1991); Campbell and Ramey (1994, 1995).

[g]Ramey et al. (1985); Wasik et al. (1990).

[h]IHDP (1990); McCormick et al. (1991, 1993); Ramey et al. (1992); Brooks-Gunn et al. (1994a, b); McCarton et al. (1997).

[i]Olds et al. (1986a, 1986b, 1988, 1997); Olds, Henderson, and Kitzman (1994); Olds (1996).

A second feature of program design concerns which members of the family receive program services and during what period of child development. Among the programs summarized in Table 2.1, some concentrate resources on treating the parents—or more typically the mother—because changing the mother's behavior may improve the child's development (e.g., through improved emotional support, better cognitive stimulation, or an increase in economic self-sufficiency). Interventions that focus on the mother often begin during the prenatal period and continue through the early period of child development. Other programs focus on treating the child, often intervening during the period just before entering school—ages 3 and 4. As noted in Table 2.1, some programs combine these two approaches and treat both parents and children during different stages of early childhood.[5]

Third, the programs we review in Table 2.1 also vary in the location and nature of the services provided. Some programs offer services in the family's home, others provide services in a center setting—typically a child care center—and others may provide services in both. Guided by program objectives, the types of services often vary by whether they are provided to the mother or child, and whether they are provided in the home or a center setting. For instance, a program designed to improve child health might provide the mother training in parental skills in the home; for the child, there might be home-safety inspections and child-abuse recognition in the home and health screenings in a clinic. Another program that strives to promote school readiness might only provide services to the child in a center—such as social interaction and cognitive stimulation. Programs providing services for children in centers may also have the added objective of facilitating the mother's employment because they provide child care.

---

[5]Another model, called "two-generation" programs, adopts a three-pronged approach by providing adult-focused, parent-focused, and child-focused services. In these programs, family support and preschool education activities are combined with parent education and job training, the latter aimed at increasing economic self-sufficiency. Examples of large-scale programs include Even Start Family Literacy Program, New Chance, the Comprehensive Child Development Program, and the JOBS Child Outcomes Study. Smaller-scale examples include the Avance Family Support and Education Program. For recent reviews of these and other two-generation programs, see St. Pierre, Layzar, and Barnes (1995) and IRP (1997b).

Finally, as summarized in Tables 2.2 and 2.3, the benefits from targeted early intervention programs have been measured for children and their parents (largely for mothers) in four domains: emotional and cognitive development, education, economic well-being, and health.[6] As we highlight in the discussion that follows, gains in the child's cognitive or emotional development have been measured through improved behavior and higher IQ in the treatment versus the control group, although these gains typically erode soon after the intervention ends. Educational improvements have been observed for the child in several areas, including better achievement test scores, increased rates of school completion, faster promotion from grade to grade, and reduced participation in special education programs. Improvements in economic well-being have been measured for children, although these behavioral differences are not observed until later follow-ups when the child reaches adolescence or young adulthood. Finally, interventions have also produced health benefits for children, including reduced levels of child abuse and emergency room visits.

Our understanding of the effects of early intervention programs on a parent's behavior is more limited, since only a few of the studies we review include any assessment of parental outcome differences and then almost exclusively for the child's mother. Favorable benefits for mothers include more satisfactory parent-child relationships, improved educational attainment and labor force participation, reductions in welfare utilization and criminal behavior, and better health outcomes.

In the remainder of this chapter, we turn to a more detailed discussion of each of the ten programs summarized in Tables 2.1–2.3. Most of the programs we highlight were carried out in the 1960s and 1970s; a few began in the 1980s. Our discussion places these efforts in their

[6]As we discuss below, the numerous evaluations of Project Head Start are not readily summarized in the format used in Tables 2.2 and 2.3 (see Barnett, 1995, for a recent review of many of the Head Start evaluations). Thus, while we include Head Start in our review of historically important intervention programs, we have elected not to include Head Start in the tables that summarize program benefits. In addition, some of the studies we review collected information at follow-ups for other outcomes, including various attitudinal measures for children and parents. See, for example, those summarized for the Early Training Project (Gray, Ramsey, and Klaus, 1982) and the Perry Preschool Program (Berrueta-Clement et al., 1984; Schweinhart et al., 1993).

Table 2.2

Children's Measured Outcomes and Results for Selected Targeted Early Intervention Programs

Program	Cognitive and Emotional Development	Education	Economic Well-Being	Health
Early Training Project	IQ at age 6 (W): E>C E=95.0, C=82.8 IQ at age 7 (W): E=C E=97.7, C=91.3 IQ at age 17 (WR): E=C E=78.7, C=76.4	Achievement tests at age 7: E > C Achievement tests at age 10: E = C Achievement tests at age 17: E = C Special education by age 18: E<C E=3%, C=29% Grade retention by age 18: E=C E=58%, C=61% High school graduation rate by age 18: E=C E=68%, C=52% Complete high school after pregnancy by age 18: E > C E=88%, C=17%	NM	Teen pregnancy through age 18: E=C

Table 2.2—continued

Program	Cognitive and Emotional Development	Education	Economic Well-Being	Health
High/Scope Perry Preschool Project	IQ at age 5 (SB): E>C E=94.9, C=83.5 IQ at age 7 (SB): E>C E=91.7, C=87.1 Q at age 8 (SB): E=C E=88.1, C=86.9 IQ at age 14 (W): E=C E=81.0, C=80.7	Achievement tests at age 9: E>C Achievement tests at age 14: E>C High school GPA at age 19: E > C Time in special education through age 19 (% of years): E<C E=16%, C=28% Years in educable mentally impaired programs through age 27: E<C E=1.1, C=2.8 Years retained in grade through age 27: E=C E=0.5, C=0.7 High school graduation rate by age 27: E > C E=66%, C=45% Postsecondary education credits by age 27: E=C E=33%, C=28%	Ever arrested by age 27: E<C E=57%, C=69% Lifetime arrests through age 27: E<C E=2.3, C=4.6 Employment rate at age 19: E>C E=50%, C=32% Employment rate at age 27: E=C E=71%, C=59% Monthly earnings at age 27 (1993 $): E>C E=$1219, C=$766 Received public assistance at age 27: E<C E=15%, C=32% Received public assistance in last 10 years at age 27: E<C E=59%, C=80%	Teen pregnancies per 100 females through age 19: E=C E=68 C=117
Chicago CPC	Teacher ratings of school adjustment at age 9: E=C	Achievement tests at age 9: E > C Achievement tests at age 14: E>C Special education by age 9: E=C E=8%, C=9% Special education (yrs) through age 14: E<C E=0.6, C=0.9 Grade retention by age 9: E < C E=19%, C=26% Grade retention by age 14: E<C E=25%, C=37%	Delinquency rate through ages 13–14: E<C Delinquency rate through ages 15–16: E=C	NM

Table 2.2—continued

Program	Cognitive and Emotional Development	Education	Economic Well-Being	Health
Houston PCDC	IQ (B) at age 2: E>C E=98.8, C=90.8 IQ (SB) at age 3: E=C E = 108.1, C=104.0 Behavior problems at ages 4–7: E<C	Achievement tests at ages 8–11: E>C Grades at ages 8–11: E=C Special education by ages 8–11: E=C E=14%, C=17% Grade retention by ages 8–11: E=C E=16%, C=23% Bilingual education by ages 8–11: E<C E=14%, C=36%	NM	NM
Syracuse FDRP	IQ (SB) at age 3: E>C E=110.3, C=90.6 IQ (SB) at age 6: E = C E=109.1, C=105.0 Socioemotional behavior at age 3: E>C Socioemotional behavior at age 6: E<C	Grades at age 15: E > C (girls only) Special education by age 15: E = C Grade retention by age 15: E = C Attendance at age 15: E > C (girls only) Teacher ratings at age 15: E > C (girls only)	Referred to probation by age 15: E<C E=6%, C=22%	NM
Carolina Abecedarian	IQ at age 5 (WP): E > C E=101.4, C=93.7 IQ at age 8 (WR): E > C EE+EC=97.8, CE+CC=93.3 IQ at age 12 (WR): E > C EE+EC=93.7, CE+CC=88.4 IQ at age 15 (WR): E = C EE+EC=95.0, CE+CC=90.3	Achievement tests at age 8: E > C Achievement tests at age 15: E > C Special education by age 15: E < C EE+EC=25%, CE+CC=48% Grade retention by age 15: E < C EE+EC=31%, CE+CC=55%	NM	NM

Table 2.2—continued

Program	Cognitive and Emotional Development	Education	Economic Well-Being	Health
Project CARE	IQ at age 1 (B): E1 > E2, C E1=119.3, E2=107.8, C=108.5 IQ at age 3 (SB): E1 > E2, C E1=104.5, E2=88.4, C=92.9 IQ at age 4.5 (M): E1 > E2, E1 = C E1=103.1, E2=89.9, C=96.0	NM	NM	NM
IHDP[a]	IQ at age 3 (SB): E > C E=93.6, C=84.2 IQ at age 5 (WP): E > C (HLBW only) E=95.4, C=91.7 IQ at age 8 (W3): E > C (HLBW only) E=96.5, C=92.1 Behavior problems at age 3: E<C Behavior problems at age 5: E<C (HLBW only) Behavior problems at age 8: E=C	Achievement tests at age 8 (math only): E>C (HLBW only) Grade repetition by age 8: E=C E=14%, C=15% Special education by age 8: E=C E=17%, C=20%	NM	General health at age 8: E=C

Table 2.2—continued

Program	Cognitive and Emotional Development	Education	Economic Well-Being	Health
Elmira PEIP[b]	IQ at age 3 (SB): E=C E=103.6, C=102.0 IQ at age 4 (SB): E=C E=111.5, C=108.9	NM	Arrests through age 15: E<C (HR only) E=0.24, C=0.53 Convictions through age 15: E=C E=0.13, C=0.18	ER visits, ages 25–50 months: E<C E=1.0, C=1.5 Hospital days, ages 25–50 months: E>C E=0.54, C=0.30

NOTES: Age references are to the age of the focal child. Results that are not statistically significant are designated by E = C; results that are significant at the 0.05 level or better are designated by E > C or E < C.

NM = not measured; E = experimental, C = control; for Carolina Abecedarian, the first letter indicates preschool treatment and the second school-age, e.g., EC = experimental during preschool, control at school age. For IQ tests: B = Bayley Mental Development Index; M = McCarthy Scales of Children's Abilities; SB = Stanford-Binet; W = Wechsler Intelligence Scale for Children; WP = Wechsler Preschool and Primary Scale of Intelligence; WR = Wechsler Intelligence Scale for Children–Revised; and W3 = Wechsler Intelligence Scale for Children–III.

[a]Results are for full sample unless otherwise indicated. Heavier low-birthweight (HLBW) infants are those born greater than 2,000 grams.

[b]Results are for full sample unless otherwise indicated. The higher-risk (HR) sample is defined as single mothers with low SES for child results through age 15.

Table 2.3

Parent's (Mother's) Measured Outcomes and Results for Selected Targeted Early Intervention Programs

Program	Emotional/Parenting Development	Education	Economic Well-Being	Health
Early Training Project	NM	NM	NM	NM
High/Scope Perry Preschool Project	NM	NM	NM	NM
Chicago CPC	Parental involvement in school at age 9: E > C	NM	NM	NM
Houston PCDC	Mother-child interactions at age 3: E>C HOME Inventory at age 3: E>C	NM	NM	NM
Syracuse FDRP	NM	Completed high school by age 5: E>C	NM	NM
Carolina Abecedarian	HOME Inventory at age 4-1/2: E=C	Years of education at age 4-1/2: E>C E=11.9, C=10.3	Employment/occupational status at age 4-1/2: E>C	NM
Project CARE	Childrearing attitudes at age 3: E1=E2=C HOME Inventory at age 4-1/2: E1=E2=C	NM	NM	NM
IHDP	Mother-child interactions at age 2-1/2: E>C HOME Inventory at age 3: E>C	Months in school through age 3: E = C E = 4.9, C=4.2	Months employed through age 3: E>C E=16.7, C=15.6 Months on public assistance through age 3: E = C E=14.4, C=12.6	Subsequent pregnancy through age 3: E=C

Table 2.3—continued

Program	Emotional/Parenting Development	Education	Economic Well-Being	Health
Elmira PEIP[a]	HOME Inventory at 46 months: E=C Reports of child abuse and neglect through age 15: E<C E=0.29, C=0.54	Years of education at age 4: E=C E=11.4, C=11.1	Months employed through age 15: E=C E=96.4, C=89.7 Months on AFDC through age 15: E<C (HR only) E=60.4, C=90.3 Months on Food Stamps through age 15: E<C (HR only) E=46.7, C=83.5 Months on Medicaid through age 15: E=C E=61.8, C=70.0 Arrests through age 15: E<C (HR only) E=0.18, C=0.58 Convictions through age 15: E<C (HR only) E=0.06, C=0.28 Jail days through age 15: E<C (HR only) E=0.04, C=1.11	Subsequent pregnancies through age 15: E < C (HR only) E=1.5, C=2.2 Subsequent births through age 15: E < C (HR only) E=1.1, C=1.6 Months between first and second birth through age 15: E > C (HR only) E=64.8, C=37.3 Substance use impairments through age 15: E<C (HR only) E=0.41, C=0.73

NOTES: Age references are to the age of the focal child. Results that are not statistically significant are designated by E = C; results that are significant at the 0.05 level or better are designed by E > C or E < C. NM = not measured; E = experimental; C = control. The HOME Inventory assesses aspects of parental care giving and characteristics of the physical home environment.

[a]Results are for full sample unless otherwise indicated. The higher-risk (HR) sample is defined as single mothers with low SES for mother results through the child's age 15.

historical context. This presentation of historically important interventions and their evaluations provides a basis for understanding the current policy debate regarding the future of investments in early childhood programs. Early intervention efforts today clearly echo the debates of prior eras. Reflection on the prejudices and limitations of earlier work may improve the outcomes of future efforts.

First Models for Targeted Early Intervention

Many historical accounts have been written about the early intervention programs that debuted in the 1960s (Zigler and Muenchow, 1992; Zigler and Styfco, 1993; Richmond and Ayoub, 1993). Here we discuss three of the most influential.

Early Training Project

Early intervention theory and practice today were shaped by the first efforts in the early 1960s at enhancing development in children with mental retardation thought to be caused by inadequate home environments. One of the initial large studies was the Early Training Project directed by Susan Gray in Murfreesboro, Tennessee (Gray and Klaus, 1970; Gray and Ramsey, 1982; Gray, Ramsey, and Klaus, 1982; Lazar and Darlington, 1982). The motivation for this program was concern over the progressive retardation of low-income children in school. The intervention was designed to improve academic performance through better cognitive performance and achievement orientation.

The Early Training Project enrolled 65 black children from the 1958 cohort selected from a low-SES sample based on parent occupation, education, and income, and on housing conditions. Participants were randomly assigned to the intervention (either a three-summer or two-summer program) or a control group. Between 1962 and 1965, those in the treatment group received weekly home visits during the school year and participated in a 10-week part-day preschool program during the summer. Children were assessed during the intervention period and in 1965, 1966, 1968, and 1975 at approximately ages 7, 8, 10, and 17, respectively. At the time of the final follow-up in 1978, most participants were 19 years old, with 80 percent of the original participants available for long-term outcome assessment.

Early results identified a significant difference between treatment and control groups in terms of IQ, but these effects disappeared within a few years.[7] Differences in achievement tests were also significant in the first few years after the intervention: By the end of the first grade (in 1965), children receiving the intervention scored higher on three of four subtests of the Metropolitan Achievement Tests. However, by fourth grade (in 1968), the differences between the two groups of children, while in the expected direction, were no longer significant on this same test. In later follow-ups through the high school level, no significant differences were found between the experimental and control groups on a range of tests (including the WISC-R for intelligence and the Stanford TASK, a comprehensive achievement test).

Although the intervention did not produce lasting differences in IQ or achievement tests between treated and nontreated children, the evaluation did find that experimental children were less likely to ever be placed in a special education class (specifically, a class for the "educable mentally retarded"). Program participants were also less likely to be retained in grade or to have dropped out of high school, but the differences were not statistically significant. While treatment and control girls showed no significant difference in their rates of teen pregnancy, among the treatment group, those who bore children were significantly more likely to return to school and obtain their high school diploma.

High/Scope Perry Preschool Project

The High/Scope Perry Preschool Project in Ypsilanti, Michigan, provides one of the longest assessments of the effects of early childhood intervention for low-income children (Weikart, Bond, and McNeil, 1978; Schweinhart and Weikart, 1980; Berrueta-Clement et al., 1984; Schweinhart et al., 1993). Like the Early Training Project, the program was motivated by the poor school performance of economically disadvantaged children. By enrolling children in one or two

[7]Scores on the Wechsler Intelligence Scale for Children (WISC) were no longer significant at age 7 (Lazar and Darlington, 1982), while the Stanford-Binet scores remained significantly different until age 10 (Gray and Klaus, 1970).

years of preschool, the program aimed to improve cognitive and social outcomes in both the short and long run.

The initial preschool study consisted of 123 African American children and their parents who enrolled in the study over five waves between 1962 and 1967.[8] The intervention group was gathered by interviewing all families with 3-year-olds in Ypsilanti during each project year. Families were rated on the basis of parent education, occupation of head of household, and household density. For families who scored low on the rating of socioeconomic status, children were administered an intelligence test. Each participant was assessed to have an IQ (Stanford-Binet) of less than 85 on entry into the study.

Participants were randomly assigned to either the preschool program or a control group. The only exceptions to random assignment were younger siblings who were assigned to the same group as already participating older siblings, and two children who were moved to the control group because of parental conflicts with participation in the preschool. The program participants were followed annually through age 11, and again at ages 14, 15, 19, and 27, with data collected through interviews, school records, and public record reviews. At the age 27 follow-up, 117 of the 121 living participants completed interviews. Attrition of the study groups averaged about 9 percent and was not significantly different between the treatment and control groups, nor was attrition related to characteristics at program entry (Barnett, 1993). The long follow-up for this study and the low rate of attrition sets it apart from most others in the literature. The relatively small sample sizes, however, are a drawback.

Over the course of the intervention, 45 children entered the Perry Preschool at age 3 and attended for two years, while 13 entered at age 4 and attended one year before entering public kindergarten. The study intervention consisted of daily 2-1/2 hour center-based classes and weekly 90-minute teacher home visits, both delivered from October to May each year. The in-class teacher-student ratio was 1 to 6, and all the teachers were certified public school teachers trained in child development. The curriculum evolved over the five years of the

[8]Five children in the original sample of 128 children were lost because of migration or death during the preschool treatment period.

study from a nursery school to a Piagetian-based early education approach (Barnett, 1993).

The study was well designed to provide information on a broad range of outcomes, including cognitive development, academic achievement, work activities, welfare participation, and criminal records (Schweinhart et al., 1993). Despite the small sample sizes, both short- and long-run assessments demonstrate favorable effects for children participating in the preschool program across a broad spectrum of outcomes (see Table 2.2).[9] As with other early intervention studies of the era, the first outcomes measured were changes in IQ. At the end of the program intervention, children in the preschool program had IQ scores that exceeded the control group by over 11 points. The favorable IQ effect for program participants began to decline after school entry, disappearing by second grade (age 8) (Schweinhart and Weikart, 1980).

These early favorable IQ effects were followed by improved academic achievement even after differences in IQ between the groups ceased to be statistically significant. For instance, achievement test scores for program participants remained significantly higher than those for the control group through age 14. Preschool participants had better grades and were more likely to have graduated from high school; at age 28, however, there were no differences in postsecondary education participation (Schweinhart et al., 1993). Time in special education was significantly lower for program children at ages 19 and 27. The difference in grade retention, while in the expected direction, was not significant at age 27.

At the last follow-up, other lasting differences were evident as well in employment, welfare, and crime outcomes (Schweinhart et al., 1993; Barnett, 1993). For instance, by age 27, program participants had significantly lower rates of current and past welfare utilization (i.e., AFDC—Aid to Families with Dependent Children, Food Stamps, general assistance, and so on). Lifetime criminal activity—both incidence and severity—was also significantly lower. Employment rates and earnings for program participants were correspondingly higher,

[9]Results have been reported for other measures in addition to those summarized in Table 2.2, including other economic and social behaviors and attitudes (see, e.g., Berrueta-Clement et al., 1984; Schweinhart et al., 1993).

although the employment rate difference was statistically significant at age 19 but not at age 27. Health effects, in contrast, were not as strong. The difference in the teen pregnancy rate by age 19 was large (64 per 100 females versus 117) but only marginally significant ($p = .08$).[10] Other life course behaviors include a statistically significant higher rate of marriage by age 27 among women participants in the preschool program (Schweinhart et al., 1993).

The Perry Preschool program, because of its design, is often closely associated with the Head Start program, an intervention model we turn to next. However, it is important to keep in mind that Perry differed in many respects from the average Head Start program as implemented today. For instance, staff ratios were lower and instructors more highly qualified and trained. At the same time, the Perry Preschool program experienced some of the same problems evident with Head Start programs. For instance, average attendance was 69 percent in the first year and home visit appointments were also missed (Barnett, 1993). Nevertheless, the study of long-term outcomes in Perry Preschool has been closely aligned with Head Start in an attempt to promote the benefits of early intervention and to draw attention to the effects of investments in program quality (Zigler and Muenchow, 1992).

Project Head Start

The most well-known of all early intervention programs, and certainly one of the largest, is Head Start. Initially designed and implemented as Project Head Start in 1965, its chief architects included child development experts Julius Richmond and Edward Zigler. The project was the result of a combination of factors, including public and policymakers' concern about the growing number of children living in poverty in the United States (Zigler and Muenchow, 1992; Richmond and Ayoub, 1993). The development of new theories of child development and the existence of a federal budget surplus also coincided with the program's initiation. Despite the limited experi-

[10]The birth rates are calculated based on the total number of pregnancies and live births per woman in treatment and control groups by age 19. The 24 women in the control group had a total of 28 births in contrast to 17 births for the 25 women in the treatment group.

ence with early intervention for socioeconomically disadvantaged children without disabilities, this large-scale program was launched.

At the time of Head Start's creation, the influence of Bloom and Hunt's theories regarding early developmental vulnerability and opportunity from environmental inputs coincided with political forces targeting government resources for altering the course of multigenerational poverty. The "cultural deficit" model of early intervention was predominant among policymakers and many social activists. It equated economic disadvantage with failure to provide the proper middle-class child-rearing experiences necessary for advancement in society (Swadener and Lubeck, 1995). Early intervention was seen as an inoculation against exposure to the detrimental influences of poverty, and it was believed to improve intelligence and lead to academic success (Lazar and Darlington, 1982; Ramey, Dorval and Baker-Ward, 1983).

Political circumstances had led to a budget surplus under the Economic Opportunity Act of 1964. Sargent Shriver, director of the Office of Economic Opportunity, was anxious to use the surplus funds for a program targeted at children (Cravens, 1993). The Kennedy family's experience with mental retardation had led them to create a foundation for research on mental retardation that included support of early education programs for economically disadvantaged children with retarded development. Shriver had been greatly impressed by a visit with his wife, Eunice Kennedy Shriver, to the Early Training Project, which the Kennedy Foundation had sponsored (Zigler and Muenchow, 1992). Shriver hoped to create a national program that would achieve the same improvements in IQ as shown by the children served by the Early Training Project.

Shriver assembled a committee of child development experts to make recommendations for a model program for low-income children. These experts included the psychologist Urie Bronfenbrenner, who had just begun to elaborate his ecological theory of child development (Richmond and Ayoub, 1993). The ecological perspective examines how a child's interactions with environment and society affect developmental outcomes. This was one of the first theories to model early childhood development not as a linear, cause-and-effect process, but rather as the confluence of many different layers of environmental factors (Meisels and Shonkoff, 1990). Bronfenbrenner's

theories were influential in the design of Head Start and subsequent early childhood interventions.

Program Implementation. As originally conceived, Project Head Start was an experimental eight-week pilot project implemented in 1965 in nearly 2,500 communities across the country, with the goal of improving the social competence of economically disadvantaged preschool children. The first trial served 500,000 4- and 5-year-old children and their families and was launched within 12 weeks of the announcement of the experimental initiative (Zigler and Muenchow, 1992). The intervention was designed to be comprehensive, with a home-visit as well as a center-based component. It provided social, health, and education services in addition to early education opportunities. Parental participation was integral to the program.

The popular and political appeal of Project Head Start, however, was based on public attention to the possibility of raising IQ scores in children rather than on its broader goals (Zigler and Muenchow, 1992). It was thought that early intervention would raise intelligence and lead to improved life outcomes for poor children (Ramey, Dorval, and Baker-Ward, 1983). The comprehensive nature of the intervention was likewise underemphasized in policy arenas at the time of its implementation and during evaluations years later.

As Zigler and Muenchow point out in their account of Head Start's implementation, there were few paradigms for early childhood programs in the United States in 1965 (Zigler and Muenchow, 1992). Only 18 states had public kindergarten programs, and there was almost no experience with preschool education. The planning document for the project had few guidelines about educational activities and requirements for training preschool staff. Head Start has been handicapped by and criticized for this lack of attention to quality of services, curriculum, and staffing; and reforms of Head Start in the 1970s began to address such quality issues.

The Head Start program has changed considerably since 1965. In its second year of funding, many programs expanded to a nine-month, half-day service, and by the early 1970s most programs were designed as school-year programs (Zigler and Muenchow, 1992). There continued to be an emphasis on local parent and community involvement in addition to national program standards. This resulted

in considerable program variation across localities (Zigler and Styfco, 1993).[11]

Many other child and family support programs were developed as a result or extension of the Head Start experience. Parent and child centers (PCCs) were launched by the Head Start Bureau in 1967 in 33 communities across the country. They were created to address the concerns that Head Start began too late in development to be effective. The PCCs were designed to serve low-income families with children up to 3 years of age through the provision of comprehensive health, social, and educational services. Project Follow Through was designed to provide continued intervention into middle childhood and addressed concerns at the other end of Project Head Start's time spectrum—when vulnerable children enter the public school system without continued support. It was thought that continued intervention in elementary school would provide long-lasting IQ effects (Zigler and Muenchow, 1992).

Other Head Start programs include Education for Parenthood, which provides parenting education for adolescents; Parent and Child Development Centers, designed to study infant development and parenting; and Child and Family Resource Programs, a service center for families with children to age 8. More recent programs include the Head Start Migrant programs, Comprehensive Child Development Centers, and Early Head Start. The latter two programs are models of comprehensive family support programs for low-income families with children under 5 years of age.[12]

Program Evaluation. The Head Start program has served over 15 million children at a cost of $31 billion since 1965 (GAO, 1997). There is a huge volume of literature that reflects the continued interest in the design and outcomes of investments in this program. An evaluation of Head Start research literature released this year by the General Accounting Office found nearly 600 citations and documents on

---

[11]Head Start providers are public and private agencies that apply for grants from the Department of Health and Human Services to cover 80 percent of program costs, with the remaining 20 percent obtained from nonfederal sources (GAO, 1997).

[12]For a recent review of these and related programs, see St. Pierre, Layzar, and Barnes (1995).

Head Start (GAO, 1997).[13] In the discussion that follows, we first focus on evaluations of Project Head Start. Because the program has been the subject of numerous other evaluations since that time, we provide a brief review of those assessments as well.

Measuring the success of Head Start is complicated by several factors. First, there has been no national randomized control trial conducted for purposes of evaluating the program as originally designed.[14] Thus, researchers have had to overcome the lack of an experimental design, which continually raises concerns over possible biases due to analytic methods. Second, the widespread implementation of the program and changes in design and program regulation over the years means that there is no single program to evaluate. Program services have varied from place to place and in response to the changing profile of the low-income children and families served.

Finally, the earliest assessments were constrained by both popular and academic attention to cognitive achievement despite the wider goals of the Head Start program. Researchers were handicapped, in part, by limited experience in the early 1960s in evaluating early childhood interventions. While methodology for studying cognitive outcomes was well developed, early evaluators had little experience or literature to guide them in the measurement of less tangible changes in child and family behaviors. This bias toward examining primarily cognitive effects continued with later assessment, as well (Currie and Thomas, 1995).

The earliest evaluations of Project Head Start, based on the program first implemented in 1965, showed some promising results, including changes in IQ as large as 10 points for the first summer program participants (Deutsch, 1967; Zigler and Muenchow, 1992). A study of a group of New York children in this first summer program assessed school readiness and found that Head Start participants were initially

---

[13]As we noted earlier, the many studies evaluating Head Start are not easily summarized in the same format as the other programs reviewed in Tables 2.2 and 2.3. Consequently, Head Start is not included among the programs summarized in those tables.

[14]Some experimental studies with randomized treatment and control groups have been conducted. However, their small samples, nonrepresentative study populations, and possible nonrandom attrition may have produced biased estimates of the effects of the overall program (Barnett, 1992, 1995).

better prepared for school, but that these effects faded within a few months (Wolff and Stein, 1966). These assessments were hampered, however, by the lack of experimental design and limited resources to measure outcomes other than IQ.

To improve upon these early assessments, a national evaluation of Head Start's first three years was commissioned by the Johnson administration. The Westinghouse Report (Cicirelli, 1969) was intended to provide—in the shortest amount of time possible (little more than a year)—a national assessment of the first three years of the Head Start effort. It used a national sample of children, with 70 percent drawn from the first summer Project Head Start participants and 30 percent from the subsequent full-year program participants. These children were matched retrospectively with a control group, and cognitive effects using standardized measures were assessed. There was also an unsuccessful attempt to assess child attitudes in school using unvalidated measures.

The Westinghouse analysis did not find any favorable cognitive effects from participation in the summer program (Cicirelli, 1969). It did find that in first and second grade, full-year participants performed better on a school readiness test, but the researchers were unable to draw conclusions about third grade performance because of the small sample size. There were also geographic and ethnic subgroup variations in performance. Based on these results, the study concluded that Head Start was not beneficial. However, the study's design, with its retrospective control group, raised serious concerns regarding the legitimacy of its conclusions. Though this design flaw could not be overcome through analytic methods, later reanalyses of this study with improved statistical modeling of the participants and control groups did find small favorable effects from the summer program, although researchers continue to debate the size of the effect (Barnow and Cain, 1977; Bentler and Woodward, 1978; Magidson and Soerbom, 1982).

The Westinghouse report coincided with a reversal in the political and societal climate of the early 1960s, which led to a shift of government priorities away from social service programs (Honig and Lally, 1982; Zigler and Muenchow, 1992). Many longitudinal early intervention studies and services such as Head Start and Project Follow

Through experienced draconian budget cuts or ceased to exist during this period (Reynolds, 1994; Honig and Lally, 1982).

At about the same time, Arthur Jensen published his monograph on genetic influences on IQ scores, emphasizing the dominance of nature over environmental influences on intelligence. He predicted that early intervention could not be successful because it is targeted at children whose potential is limited by genetics (Jensen, 1969). While his premises were controversial, the monograph served to reinforce a stigma on Head Start's first efforts at early intervention for economically disadvantaged children. Disenchantment with Head Start was also fostered by the Coleman Report, which concluded that schools were unable to overcome the socioeconomic circumstances of children participating in compensatory education programs and that the quality of early education experience had little effect on school outcomes (Coleman et al., 1966).

In 1971, Bronfenbrenner was commissioned by the Office of Child Development to review the early intervention literature to date. He also concluded that short-term benefits of program participation fade in elementary school within two to three years (Bronfenbrenner, 1974). In his report, he reemphasized the importance of family in fostering and sustaining development, and pointed to family-focused interventions as the remedy for academic "fade-out." The broader implications of his conclusions were not emphasized as much as his evaluation of the long-term effects of early intervention.

A decade later, in 1981, the Administration for Children, Youth and Families funded an in-depth review of the Head Start literature and findings called the "Head Start Evaluation, Synthesis and Utilization Project" (McKey et al., 1985). This meta-analysis reviewed over 210 reports on Head Start and 1,600 other related documents to produce a synthesis of outcomes of the first 20 years of Head Start. The final analysis was based on 76 studies that met eight broad selection criteria (McKey et al., 1985). These criteria did not include sample size, attrition, or random assignment design, and the meta-analysis is unable to account for possible intrinsic biases in the results of the original studies that are due to these factors. These limitations must be kept in mind when considering the findings of the study.

The meta-analysis, like many studies that preceded it, concluded that there were immediate cognitive benefits of participation in Head Start but that these benefits did not persist over the long term. However, based on a limited number of studies, Head Start participants were found to have better school success in terms of grade retention and special education use. Socioemotional gains were also found to be favorable but possibly transitory as well. There were definite, favorable physical health effects found.

More recent analyses have measured the effects of Head Start participation using alternative statistical methods and a broader array of outcome measures. For example, Currie and Thomas (1995) use data on children from the National Longitudinal Survey of Youth (NLSY), a nationally representative sample, to estimate the effect of Head Start on school performance, cognitive attainment, access to preventive medical care, and health and nutritional status. They find, for white children, that Head Start has significantly favorable and lasting effects on test scores and school attainment relative to participating in other preschool programs or no preschool program. Significant favorable increases in test scores for black children are also measured at young ages, but these gains eventually disappear. For both white and black children, Head Start is associated with higher immunization rates but no differences in long-run nutritional status as measured by child height-for-age.

The Next Generation of Targeted Early Intervention Programs

Early intervention models in the 1970s evolved as researchers and service providers became more cognizant of the role of family, culture, and society in child outcomes. The conclusions of the first Head Start studies that emphasized the overwhelming effect of the parents and home on an intervention's effectiveness helped shape the family intervention models that followed in the 1970s. Bronfenbrenner's ecological model of development was well known, and there was a better understanding of the interaction of the child with family and environment. Consequently, many of the model programs in this period were designed to intervene even earlier in childhood, with parents as a focal point as well as the child. Interventions and their evaluations were designed with an eye to disproving

Jensen's theory and the negative conclusions of the Westinghouse study (Honig and Lally, 1982). Below we discuss five of the most prominent early intervention programs initiated during the 1970s.

Chicago Child-Parent Center and Expansion Program

The Chicago Child-Parent Center (CPC) was started in 1967 in 11 Chicago public schools to serve economically disadvantaged children age 3 to 5 years (Reynolds, 1994, 1997; Reynolds and Temple, 1995; Reynolds, Chang, and Temple, 1997). With federal funds (through Title I, now Chapter I), the preschool program initially provided a structured half-day program during the nine-month school year for 3- to 4-year-olds, and was designed as an early education program to prepare children for school through promotion of reading and language skills. The program provided comprehensive services including health and social services and parent involvement. In 1978, with additional state funding, the program was expanded to continue services for children through third grade, including a full-day kindergarten component. Today, 24 centers provide preschool and postschool components, through grades 1, 2, or 3.

Outcomes for one recent cohort of CPC children have been followed in the Longitudinal Study of Children at Risk (LSCR), a prospective study of the adjustment to school of 1,539 low-income minority children (95 percent black, 5 percent Hispanic) in 26 Chicago-area kindergarten programs.[15] The LSCR provides an opportunity to evaluate the CPC program using a quasi-experimental design, comparing outcomes for the 1,150 participants in 20 CPCs having both preschool and after-school components with those from a control group of 389 nonparticipants in six randomly selected schools having a locally funded, full-day kindergarten program for low-income children. Children in the control group did not have access to a CPC in their neighborhood. Children in the LSCR form a single age cohort, entering the program at age 3 or 4 (1983–1985) or kindergarten (1985–1986), graduating from kindergarten in the spring of 1986. Some children continued to participate in the CPC through age 9 (1989) for a total of up to six years in the intervention. At the time of

[15]This study is also sometimes called the Chicago Longitudinal Study (Reynolds, 1997).

the latest follow-up at age 14, about 75 percent of both the treatment and the follow-up groups are available for study. Data on the children's outcomes include achievement tests from school records, and interviews with teachers, parents, and children to assess scholastic, social, and psychological development.

Unlike Head Start, the CPCs are integrated with the primary schools: The preschool and kindergarten programs meet in buildings of their affiliated elementary schools. While the curriculum is not uniform across centers, they all aim to provide comprehensive health, social, academic, and school support services to promote school readiness, with a special emphasis on reading and language skills and parental involvement. During the preschool years, children attend a structured half-day program during the nine-month school year; the kindergarten consists of a full-day (six hour) school-year program. The primary grades program provides reduced class sizes, parental involvement activities, and instructional coordination. Other comprehensive services at each age include free breakfasts and lunches, and health screenings. Adult-to-child ratios average 1 to 8 and 1 to 12 for the preschool and kindergarten components, with class sizes of 17 and 25, respectively. Average class sizes for the control children in kindergarten and primary grades was 30.

Examining data covering the preschool, kindergarten, and primary grades, researchers have used psychometric and econometric techniques designed to account for the nonrandom participation of children in the CPC—through controls for relevant background variables and explicit modeling of the selection mechanism (Reynolds, 1994, 1997; Reynolds and Temple, 1995). Because children entered the program at different ages and participated for varying lengths of time in both preschool and postschool intervention, it is possible to examine the effects of the timing (i.e., entry point) and duration of treatment on outcomes, albeit without the advantage of a randomized control design.

At the end of the intervention at age 9, those who participated in the CPC had significantly higher reading and math achievement scores, lower rates of grade retention and higher ratings of parental involvement. There were no significant differences found, on average between participants and nonparticipants, in special education placement and teachers ratings of school adjustment at age 9, al-

though years of special education were significantly lower for treatment children by age 14 (Reynolds, 1994, 1997). The differences in achievement scores between groups tended to become smaller over time although they remained significant through age 14 for math scores. The findings for regression-controlled mean differences are generally similar to those based on models that explicitly consider selective program participation (Reynolds and Temple, 1995).

A comparison, based on time in the intervention, across children showed strong duration effects: At age 14, six-year program participants had the largest gains compared with nonparticipants in terms of reading and math scores, grade retention, and years of special education (Reynolds, 1997). For most of the outcomes studied through age 9, the association between time in the program and program benefits, rather than being linear, exhibited a threshold effect: Those who participated four years or more had the greatest benefits (Reynolds, 1994). Moreover, the results indicate that participation in both preschool and primary grade components confers the greatest benefits. Results based on data through age 14 suggest that, controlling for duration in the program, there were no significant differences between those who began with the preschool versus the kindergarten component.

Finally, researchers have recently examined measures of problem, illicit, or illegal behavior in grades 7 to 10 (Reynolds, Chang, and Temple, 1997). While differences in delinquency rates between treatment and control groups based on time in the program were significant at ages 13 to 14, these differences were no longer evident at ages 15 to 16.

Houston Parent-Child Development Center

The Parent-Child Development Centers (PCDCs) were an outgrowth of the CPCs. The PCDCs were among the first attempts by the Office of Economic Opportunity (OEO) to evaluate models of improving child outcomes through family intervention customized to individually assessed needs. The Houston PCDC was one of OEO's model research and demonstration projects designed to promote parental change by educating a sample of low-income Mexican-American

mothers as well as their children (Johnson et al., 1974; Andrews et al., 1982; Johnson and Breckenridge, 1982; Johnson and Walker, 1991).[16] To address the concern that Head Start services started too late, the Houston PCDC provided services beginning when children were 1 year old.

Program participants were recruited from Houston neighborhoods with the lowest levels of family income and adult education, and high concentrations of Mexican Americans. Mexican American families with a child age 1 year old were randomly assigned to treatment or control groups. Families were recruited across eight cohorts to participate in the program between 1970 and 1980. To date, results are available for the first five cohorts through a follow-up when the children were between the ages of 8 and 11 (Johnson and Walker, 1991). Attrition rates have been very high, with only about 50 percent of the original sample available for follow-up. A comparison of sample characteristics at the latest follow-up shows no evidence of bias due to selective dropout or follow-up attrition, but the high attrition rate is problematic nonetheless. Preliminary results are available for a more recent follow-up, when the oldest children were age 16 (Institute for Research on Poverty, 1997a).

In the first year of the program, a paraprofessional educator made an average of 25 home visits designed to provide the mother with improved skills in teaching her infant, especially in the area of language development. In the second year, mothers and children were treated in a center setting for up to four half-day sessions per week. At the center, mothers participated in a curriculum that emphasized child cognitive and language development and other aspects of child rearing, health, and safety. Children were in a Piagetian-based nursery school while their mothers attended these classes. English language classes were also available to mothers on a voluntary basis. Fathers were encouraged to participate as well through weekend sessions during the first and second years.

---

[16]In the initial phase, from 1970 to 1975, the PCDC program was also implemented in Birmingham, Alabama, and New Orleans, Louisiana. For initial results of these programs, which differed to some extent from the Houston model, see Andrews et al. (1982).

At the earliest follow-up, children in the program (first two cohorts) were found to have higher IQ scores. But the difference, which was significant at age 2, was only marginally significant ($p < .10$) by age 3 (Andrews et al., 1982). Mother-child interactions and the home environment, predictors of later behavior problems, were better in the treatment group at age 3. Behavior problems, which were measured at ages 4–7 (first four cohorts), were also significantly better for the treatment group (Johnson and Breckenridge, 1982). By the time the children were ages 8 to 11, those who participated in the intervention had a decreased need for bilingual education services, with 14 percent of experimental children versus 36 percent of control children enrolled in bilingual education at the time of the follow-up survey (Johnson and Walker, 1991). Although achievement test scores were significantly higher for the treatment group at the same age, there were no significant differences in grade retention, special education referral, or school grades. Preliminary results at ages 9–16 show no differences in delinquency rates for program participants versus controls (IRP, 1997a).

Syracuse Family Development Research Program

The Family Development Research Program (FDRP) was implemented from 1969 to 1975 in Syracuse, New York. It consisted of adding intensive, comprehensive family services, including home visits by paraprofessionals starting prenatally, to the already established early education and child care program at Syracuse University (Honig and Lally, 1982; Lally, Mangione, and Honig, 1988). The intention of the intervention was to improve the quality of life for the children and their mothers. It was hoped that these adjunctive services would give rise to better parenting and home environment—and improved parental autonomy and self-sufficiency—and would lead to improved outcomes for the participating children.

At the outset of the FDRP, 108 families, mostly African American, were enrolled. These families had low income (less than $5,000 a year in 1970 dollars), and the family head had less than a high school education and only semi-skilled employment experience, if any. Eighty-five percent of mothers were single parents, and the mean age was 18 years. The design of the study was not randomized. A control

group matched for age and sex with the FDRP children was recruited three years into the intervention.

Evaluation of the intervention was limited by loss of funding for long-term follow-up after the children completed second grade. Funding was eventually obtained for a follow-up that focused on collecting information for the treatment and control groups through interviews with the child, parents, and teachers, and through records from schools and various components of the criminal justice system. Of the original 108 families in the intervention, 82 completed the full program, and 65 (60 percent of the original enrolled sample) consented to follow-up at 14 years. Of the control group, 74 completed the study and 54 (50 percent of the original control group) provided consent for the 14-year follow-up. Interviews with parents and children were actually completed for a smaller sample in each group (51 and 42 parents and 49 and 39 children in treatment and controls, respectively). Although the attrition rates are high, the follow-up treatment and control samples did not differ significantly on a series of observed indicators from the group that completed the program. Nevertheless, sources of potential bias include nonrandom attrition associated with unobserved factors and the delayed matched-control design.

Families participating in the intervention received information about child nutrition, health and safety, and social services, starting in the third trimester of pregnancy and continuing until the intervention children reached 5 years of age and entered school. Weekly home visits provided an ongoing assessment of service needs and referrals in the community as well as parenting and child development education. Parents in the intervention participated in the program design and delivery through a formal parent organization similar to that of the Head Start model. The intervention was run in conjunction with the Syracuse University Children's Center preschool program that offered year-round, quality day care; the center provided a specialized curriculum of early education for the infants and children enrolled in the intervention. From 6 to 15 months, infants were cared for in a half-day day-care setting, with a ratio of four infants to one caregiver. Preschoolers from 15 to 60 months attended a full-day, family-style day-care program that emphasized development of social as well as cognitive skills through an unstructured learning environment.

In the short term, the study found differences in IQ and language skills between the groups, with an IQ advantage at age 3 nearly equal to 20 points. While the treatment group still maintained a significant IQ advantage at age 4, by age 6 the difference was small and no longer significant (Honig and Lally, 1982). Socioemotional behavior in the intervention group was observed to be superior at 5 years of age but was actually inferior during first grade. It was postulated that the public school experience produced this effect because of its contrast with the quality preschool educational experience (Lally, Mangione, and Honig, 1988). Positive effects for mothers were also reported, with more intervention mothers than control mothers completing high school during their five-year participation (Honig and Lally, 1982).

In the age-15 follow-up, the study found that there was a statistically significant advantage in school attendance and grades among girls who completed the program, relative to those in the control group. Boys in the intervention did not score higher on these measures than those in the control group, and there were no differences in other educational measures between intervention and control groups for either gender. Rates and severity of juvenile delinquency, as measured by Probation Department records, were significantly lower for the treatment group compared with controls (Lally, Mangione, and Honig, 1988).

Carolina Abecedarian

The Carolina Abecedarian project was started in 1972 at the Frank Porter Graham Child Development Center of the University of North Carolina (Ramey, Dorval, and Baker-Ward, 1983; Ramey and Campbell, 1984, 1991; Campbell and Ramey, 1994, 1995). The goal of the intervention was to prevent mild mental retardation and improve academic and social competence at school entry for economically disadvantaged children. Specifically, the study was designed to examine the relative effects of day-care early education and other methods of early intervention on intellectual functioning and early academic achievement among the most disadvantaged families. In contrast to the other studies we have reviewed thus far, the Abecedarian program began just weeks after the child's birth, with a high-quality educational day-care program.

Families were recruited between 1972 and 1977 from prenatal clinics and social service agencies for participation on the basis of their scores on a "High-Risk Index" that examined demographic and psychosocial factors thought to contribute to school failure. Other criteria in the index included level of parental education, income, intelligence, and antisocial and maladaptive behaviors in the family. While 122 families identified met the criteria, the final sample consisted of 109 families (because of refusals and the exclusion of several children for mental retardation and nonrandom reassignment to one of the treatment groups); 111 children from those families were enrolled in the study. The resulting sample was 98 percent African American, with mothers who were age 20 on average, and mostly single, with a first-born child. The children were randomly assigned by 6 weeks of age to either the preschool intervention or a control group. At age 5, at the time of kindergarten entry, all the children were reassigned to either school-age intervention through age 8 or a control group. Of the original 111 children, 92 were followed until age 15.

The day-care/preschool educational program was a full-day, year-round, center-based intervention with an infant/toddler-to-teacher ratio of 3 to 1, and a child-to-teacher ratio of 6 to 1. Special curricula were developed for each age group, including a transitional program for the children entering the kindergarten intervention group. Children also received medical services at the center. The intervention program during the primary grades consisted of providing the children with extra exposure to academic concepts by having their parents engage in specific supplemental educational activities at home. A home/school resource teacher (HST) provided parents with curriculum material and other support about every two weeks (13 to 15 times a year on average), in addition to other parental support and advocacy with school staff. On alternating weeks (14 to 18 times a year on average), the HST visited the child's classroom so that materials could be customized based on the school curriculum and the child's progress.

At the end of the preschool intervention, the treatment group significantly outperformed the controls in terms of IQ, with a 7 point difference in the Wechsler Preschool and Primary Scale of Intelligence (Ramey and Campbell, 1991). IQ scores at 8 years of age continued to be significantly higher for the preschool participants than for the other children (Ramey and Campbell, 1991). Most notable, four

years after the intervention at age 12, this favorable and significant difference was also found; but the difference, while favorable, was no longer significant by age 15 (Campbell and Ramey, 1994, 1995). The study design, with treatment groups that experienced only the preschool intervention, only the postschool intervention or both, demonstrated that the strong and longer-lasting IQ effects derived from the preschool component of the intervention; in other words, the postschool intervention had little effect on the IQ differences observed between treatment and control samples.

Although the IQ effects diminished with time, at age 15, children who had participated in the preschool program still had significantly higher scores on tests of reading and mathematics, effects that had also been observed at ages 8 and 12 (Campbell and Ramey, 1995). They also had less grade retention (significantly different at ages 8, 12, and 15) and less special education placement (significantly different at age 15 only). The persistent achievement test and education outcome results, in contrast to IQ, were attributable in part to the intervention that continued after primary school entry. These results are suggestive of a role for high-quality, educational, center-based care in promoting early cognitive development, combined with subsequent school-age intervention to boost academic performance.

Finally, maternal interviews when the child was 54 months old showed that treatment mothers had significantly more years of education (after having had no difference at the start of the intervention) (Ramey, Dorval, and Baker-Ward, 1983). In addition, mothers in the program were also less likely to be unemployed and more likely to have a skilled or semi-skilled job.

Project CARE

Project CARE (Carolina Approach to Responsive Education), also at the University of North Carolina, was designed as a follow-up study to the Abecedarian Project (Ramey et al., 1985; Wasik et al., 1990). This time, home-based early intervention was tested in addition to the center-based, educational, day-care program examined in the original study. The experimental design was structured to allow comparisons between more and less intensive early educational treatment programs.

Starting in 1978, 65 families participated in the study and were randomly assigned to intervention services that lasted five years. Families were recruited from local hospital records of births over an 18-month period and screened using the High-Risk Index. Over 90 percent of the sample children were black, and about three out of four were from single-parent households. Mothers on average were age 22, and 60 percent were first-time mothers. All but one of the families agreed to participate in random assignment to (1) home visiting, (2) center-based intervention and home visiting, or (3) the control group. Families in the intervention groups began receiving home visits one month after the child's birth, and the center-based intervention began between 6 weeks and 3 months of age. One death and the relocation of four families reduced the sample size to 59 families by the end of the study, when the children were age 54 months, an average attrition rate of just under 10 percent.

Like the Abecedarian program, the center-based program consisted of full-day, year-round, educational day care, with low staff-to-child ratios and a well-structured curriculum. The family education component involved home visits by a trained family educator (usually the day-care teacher for children in the center-based program) about two and one-half times per month for the first three years and less frequently in years four and five (slightly more than one visit per month on average). Most visits (over 90 percent) were with the mother and lasted about an hour. Through child learning activities, the educators focused on child development, as well as on training in family problem-solving and parental skills. Parents in both intervention groups received additional child-rearing support and information about community resources through monthly workshops. All treatment and control children benefited from free iron-fortified formula through 15 months of age, as well as access to free medical care and other social services in the Chapel Hill area.

Children were assessed at 6- to 12-month intervals for five years using standardized tests of development, intelligence, and home environment and parenting. The strongest cognitive outcomes were found in the children who received the center-based services as well as home visiting. While the three groups showed no difference in cognitive tests at 6 months, by 1 year the children who received both day care and home visits had significantly higher scores than either the other treatment group (home visits only) or the controls (Ramey

et al., 1985). The advantage of this group over the other two grew smaller as the children approached 5 years of age, and it is theorized that this was a benefit of the participation of the home-visit only and control groups in community day care (Wasik et al., 1990). The group receiving family education alone never showed significant differences in cognitive outcomes compared with the control group. The three groups did not differ over the course of the study in the degree to which the home environment was judged conducive to development or in terms of child-rearing attitudes (Wasik et al., 1990).

More Recent Models of Targeted Early Intervention

Guided by the results of the model programs developed in the 1970s, and the ongoing evaluations of even earlier interventions, researchers have continued to design and evaluate alternative targeted intervention programs. More recent efforts have used alternative criteria to identify families and children at risk. However, the focus on intervention at the earliest ages continued, with programs that started with the birth of the child, if not earlier. Evaluations were more carefully designed to incorporate larger samples in both experimental and control groups (which permits stratifying results by factors thought to be associated with risk), and long-term follow-up. In addition, more recent studies examine an expanded range of outcomes for both participating children and their families. Two of the better-designed recent interventions are discussed here.

Infant Health and Development Project

The Infant Health and Development Project (IHDP) built upon the work of the Abecedarian and Project CARE studies but targeted infants who were born prematurely (less than 37 weeks gestation) and with low birthweight (less than 2,500 grams) (IHDP, 1990; McCormick et al., 1991, 1993; Ramey et al., 1992; Brooks-Gunn et al., 1994a, 1994b; McCarton et al., 1997). Such children are considered at risk for long-term intellectual difficulties and delayed development. The focus on biological risk, which may be strongly associated with socioeconomic risk, sets this program apart from the others we have discussed.

The study was a multisite randomized clinical trial designed to assess the efficacy of educational and family support services in preventing cognitive deficits in this population. Infants were enrolled in the study during a nine-month period in 1985 and were selected from infants born at eight participating medical institutions. Among the 4,551 infants identified as eligible based on birthweight and gestation, 3,249 were excluded for the following reasons: home too great a distance from the intervention site, physical exam consistent with greater than 37 weeks of gestation, or discharge before or after the designated recruitment period. Of the remaining group, 274 parents refused to participate, and 43 withdrew after randomization. The final sample size was 985 infants in the treatment and control groups with a sample that was about 50 percent black and 10 percent Hispanic. At the time of the last follow-up at age 8, 874 children were available for follow-up, about 90 percent in each of the groups.

The intervention began after discharge from the hospital and continued until the infant reached 36 months (corrected for prematurity). All children in both the intervention and control groups received medical, developmental, and social assessments, with referral for further care as indicated. The intervention participants also received (1) home visits weekly for the first year and then biweekly for an average of 67 visits over three years; (2) a full-day, year-round, center-based, educational day-care program starting at 12 months of age with average attendance of 267 days per year over both years; and (3) a series of parent group meetings every other month in years two and three. The primary outcome measures were cognitive development, behavioral competence ranked by the mothers, and health status. Long-term outcome measures at 8 years of age also included a standardized measure of academic achievement and parental report of school performance.

The outcomes of the IHDP are reported in aggregate and separately for the heavier birthweight infants (those greater than 2,000 grams) and the lighter birthweight infants (those less than 2,000 grams). (Tables 2.2 and 2.3 report results for the full sample, noting when results hold for only the heavier low birthweight infants.) At the end of the intervention at (adjusted) age 36 months, the participants had significantly higher IQ scores than the controls—by nearly 10 points—with the greatest difference in the heavier birthweight children. The intervention group also had higher receptive vocabulary

test scores on the developmental measure, and lower scores on maternal reports of behavior problems (Ramey et al., 1992). At 5 and 8 years of age, the significant differences in IQ were no longer apparent for the combined treatment group, although the heavier birthweight infants still had significantly better IQ scores than their matched controls (Brooks-Gunn et al., 1994a; McCarton et al., 1997). At age 8, no differences were found in rates of grade repetition or special education, or in terms of behavior problems or child health status. Math achievement scores were significantly different for the heavier birthweight group only.

Studies of the effects of the intervention on maternal outcomes have also been reported. Intervention group mother-child interactions at 30 months of infant age were judged significantly better (by a small margin) than those in the control group (Spiker, Ferguson, and Brooks-Gunn, 1993). Mothers in the intervention group were employed more months during the three years of the intervention and returned to work earlier than the control group mothers (Brooks-Gunn et al., 1994b). At the same time, there were no significant differences in welfare utilization or months in schooling over the three years, or in the rates of subsequent pregnancy.

Other analyses examined outcome differences based on participation rates and various risk factors. At 2 and 3 years of age, children in those families with the best participation in intervention services had better scores on the cognitive and development measures than those whose participation levels were lower (Blair, Ramey, and Hardin, 1995). Among poor families, the size of the effects of the program varied with the number of risk factors: larger effects for those with none to four risk factors, but no effects for those with six or more risk factors (Liaw and Brooks-Gunn, 1994). In addition to poverty, the risk factors examined included biological indicators (low birthweight, neonatal health), socioeconomic factors (race/ethnicity, parental unemployment, and mother's characteristics such as education, verbal ability, and mental health) and family structure (teenage mother, single parent).

Prenatal/Early Infancy Project

The Prenatal/Early Infancy Project (PEIP) conducted by researchers at the University of Rochester was a study of the effects of home visit-

ing on economically disadvantaged first-time mothers and their children. The study was conducted from 1978 to 1982 in Elmira, New York, and long-term follow-up on the children and their mothers is available for children at age 15 (Olds et al., 1986a, 1986b, 1988, 1997; Olds, Henderson, and Kitzman, 1994; Olds, 1996).

Five hundred first-time mothers were recruited between 1978 and 1980 from prenatal clinics and social service agencies, and 400 were enrolled in this study before their 30th week of pregnancy.[17] Women recruited for this study were those who were thought to be at high risk for poor child and family outcomes, although to avoid stigmatization, any first-time mother who asked to participate was accepted. Among the 400 participants, 85 percent had at least one of the following sociodemographic risk factors: less than age 19 at registration (48 percent), unmarried (62 percent), or low socioeconomic status (SES) (59 percent). Nearly one in four participants had all three risk factors.[18]

Participants were stratified by marital status, race, and geographic area, and were randomly assigned to one of the two intervention or two control groups. One intervention group received home visiting only during pregnancy, and the other received home visits until the children were two years old. Both intervention groups and one of the control groups received free transportation to prenatal and well-child visits. The only service that the other control group received was health and developmental screening at 1 and 2 years of age, with appropriate follow-up referrals. Differences between treatment and control group outcomes focus on the treatment group that received both prenatal and early infancy home visits (N = 116) versus the two control groups combined (N = 184).

The intervention groups were visited by registered nurses trained in parent education, methods of involving family and friends in assisting and supporting the mother, and linkage of the family with other health and human services. These services were provided to pro-

---

[17]There were no differences in age, marital status, or education between the study participants and the 100 women invited to participate but who were not enrolled.

[18]Because of the low number of minority women in this population and study (N = 46), minority participants were excluded from some analyses of maternal and child effects through age 4 of the children.

mote maternal functioning with respect to health-related behaviors during pregnancy and infancy, parental caregiving, and maternal life-course development (e.g., family planning, education, and employment). The targeted developmental processes were based on an ecological/systems model focusing on parental behaviors and social environment as important influences on pregnancy and child health outcomes. On average, nurses completed 9 visits during pregnancy (about 1 hour and 15 minutes every two weeks) and 23 visits from the child's birth to age 2.[19]

Data were collected at registration, at 32 weeks gestation, and then every four to six months for four years, with a final follow-up when the child was 15 years old. At that time, both mother and child were interviewed, and attempts were made to review state criminal and social service records for data on the mother. Reports of child abuse and neglect were also abstracted from archived state records. Of the original 300 women in the control and treatment groups described above, 245 mothers and 238 children completed assessments when the child was 15. At the final follow-up, results were reported for the full treatment plus control sample, as well as for a subsample of unmarried, low-SES mothers (samples sizes for treatment and control groups of 38 and 62, respectively).[20]

The study found significant short- and long-term advantages for both the mothers and children in the intervention group. In the short term, pregnancy behaviors were better for mothers in the intervention group, with less cigarette use, better nutrition, improved childbirth class attendance, and more social supports reported (Olds et al., 1986a). Mothers who smoked bore 75 percent fewer preterm infants than control mothers who smoked; overall, intervention group teenager mothers bore heavier infants than the control group teenagers did.

---

[19]Following birth, nurses initially visited approximately weekly up to 6 weeks of age, then every two weeks from 6 weeks to 4 months, every three weeks from 4 to 14 months, every four weeks from 14 to 20 months, and every six weeks from 20 to 24 months. The postnatal visits lasted a similar length of time (Olds et al., 1993).

[20]We note in Tables 2.2 and 2.3 when the results for the Elmira PEIP hold only for the higher-risk sample in the treatment and control groups. This occurs for child arrests and several of the mother outcomes when the child was 15, where the higher-risk sample is defined as unmarried low-SES mothers (Olds et al., 1997).

The program assessment through when the child was age 4 showed parental caregiving was also affected by participation in the intervention. There were decreased reports of child abuse and neglect during the first 2 years of life among the higher-risk families in the intervention group (Olds et al., 1986b). Fewer safety hazards and more development-promoting materials were found in the homes of the intervention group, and through age 4, these children were seen less frequently in the hospital emergency department (Olds et al., 1986b; Olds, Henderson, and Kitzman, 1994). Hospital days were significantly higher for the children through age 4 in the intervention group, although this results from one outlier in the sample that appears unrelated to the program (Olds, Henderson, and Kitzman, 1994). Through age 4 for the children in the treatment and control groups, no significant differences in IQ, completed years of education for the mother, or home environment were found (Olds et al., 1986a; Olds, Henderson, and Kitzman, 1994).[21]

The 15-year follow-up study found fewer reported acts of child abuse and neglect among the nurse-visited mothers for the full sample and the higher-risk sample (Olds et al., 1997). The other significant findings were restricted to the higher-risk sample (i.e., unmarried and low-SES mothers). For this group, months spent receiving AFDC and Food Stamps were significantly lower. The most at-risk mothers also had lower levels of criminal activity (measured by both self and state-documented data on arrests, convictions, and jail days) and reported fewer behavioral impairments due to alcohol and drugs. Although the full treatment group also spent fewer months receiving Medicaid and more months employed, the differences were not statistically significant. The beneficial effects of the program in terms of subsequent fertility continued through the 15-year follow-up, with higher-risk treatment mothers reporting fewer subsequent pregnancies and births, and a longer birth interval between the first and second child. Finally, children in the intervention group reported fewer arrests compared with those in the control group (Olds, 1996).

---

[21]For the subsample of poor, unmarried teens, the children in the treatment group, at ages 1 and 2, had IQ scores that exceeded those of the control group by about 10 points, although the difference was only marginally significant ($p < .10$) (Olds et al., 1986b).

The Elmira PEIP has been replicated for first-time mothers in a number of other communities for more diverse at-risk populations, specifically urban and minority samples. Randomized control trials began in 1990 for a sample of 1,139 low-income, unmarried, primarily African American women in Memphis, Tennessee, and in 1994 for a racially and ethnically mixed sample of 735 low-income, mostly unmarried women in Denver, Colorado. In early results for the Memphis trial, program participants demonstrated significantly lower levels of pregnancy-induced hypertension and reduced rates of subsequent pregnancies and injuries to their children through the age of 2 (Kitzman et al., 1997). Welfare utilization was also lower among program participants after two years, although the difference was only marginally significant ($p < 0.10$). There were no significant differences in children's mental development or behavioral problems, or in the mothers' education and employment. These findings track those exhibited for the Elmira sample after the first few years of the program (Olds et al., 1986a, 1986b). These samples are expected to be followed in the future so that researchers can continue to measure differences as the children mature.

SUMMARY OF FINDINGS

Given the shortcomings and limitations in the design of many early intervention evaluations and the measures omitted from them, what we don't know about the effects of early childhood intervention may exceed what we know (for more on this, see Chapter Four). Nonetheless, our review supports the proposition that, in some situations, carefully targeted early childhood interventions *can* yield measurable benefits in the short run and that some benefits persist long after the program has ended.

Moreover, as seen in Tables 2.2 and 2.3, many of the effects measured in the studies we review are sizable. For example, the Early Training Project, Perry Preschool, and IHDP found IQ differences between treatment and control groups at the end of program implementation that approached or exceeded 10 points, a large effect by most standards. The difference in rates of special education and grade retention for children at age 15 in the Abecedarian project exceeded 20 percentage points. In the Perry Preschool program, participating children had earnings that were 60 percent higher than

those of the control group when they reached age 27. There was a 33 percent difference between participating and nonparticipating mothers in the Elmira PEIP in months spent on welfare during a 15-year period. The reduction in children's crime, measured in terms of probation cases by the Syracuse FDRP, was about 70 percent. While these findings are compelling, not all of the effects measured in the studies we reviewed here or reported elsewhere in the literature are as large, and there are many statistically insignificant results (as well as a few significant results that do not conform to expectations).

Figures 2.1 and 2.2 provide one way to further summarize what we have learned from the preceding review of selected early intervention programs. For each of the studies covered in Tables 2.2 and 2.3, we note whether the evaluation measured an outcome in that domain, and if so, whether the program led to favorable and statistically significant results, to no statistically significant results, or to mixed results.[22] In particular, the white squares indicate cases in which a study did not measure the outcome designated in that column. Blue indicates that the outcome was measured and was favorable and significant. Black shows that an outcome was measured but that the program had no effect (or, rarely, had a significant unfavorable effect). Gray shading denotes those cases where results are mixed. This arises when favorable effects are found at one time but not at another (as in several studies for three of the four domains for children) and when multiple measures of crime or welfare utilization at a point in time show mixed results (as in several cases for children and mothers). Given that the ages at follow-up vary across studies and for particular measures, we also record in the blue and black squares the oldest age at which the measure is taken (which, for blue squares, also denotes the latest age at which a significant difference has been found). In the gray squares, we record only the oldest age at which a significant difference was found (there may have been subsequent measures yielding nonsignificant differences).

Given the interest in cognitive benefits from early intervention, we separately record whether a program produced IQ benefits in the short run (assessed at the end of the program intervention), and the

[22]Some outcomes reported in Tables 2.2 and 2.3 but that are measured in only one study are not shown in Figures 2.1 and 2.2.

results for longer-term IQ assessment. The long-run IQ result is sometimes mixed, with initial favorable benefits after program implementation eventually fading. This is illustrated by a gray box (where the age shown is the latest with a significant effect). Likewise, we show achievement test scores separately for short- and long-run measures.[23] Finally, the two figures show two sets of results for the IHDP (Figure 2.1 only) and the Elmira PEIP, first for the full study sample, and second for the subsample with stronger results. In the case of the IHDP, this is the heavier low-birthweight infants, a group that showed longer-lasting IQ and achievement test results. In the Elmira PEIP, the results for economic and health outcomes of mothers and crime outcomes for children are stronger for single mothers with low SES.

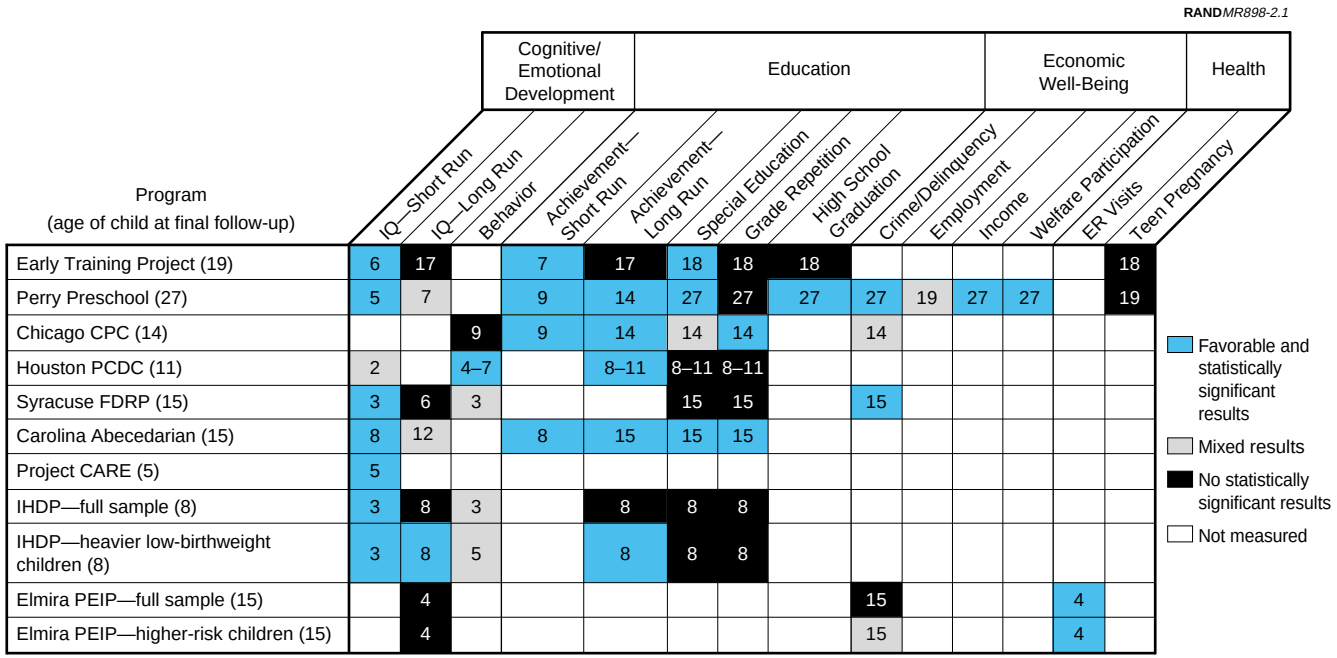
We now turn to a discussion of the general patterns that emerge from these figures, first for children and then for mothers. We link these conclusions to the larger literature on early childhood intervention.

Outcomes for Children

Most of the studies we reviewed measured effects for children, in some cases exclusively. In Figure 2.1, we first note that every study produced at least one significant benefit for the intervention children, whether in the short or long run, and most programs produced benefits in more than one domain. While we cannot attach weights to the areas that are blue or gray versus black, a simple count reveals that there is more blue and gray than there is black. Most measured effects are thus favorable.

Second, the blue areas are not limited to the first few columns. Early intervention programs affect more than just cognitive development, even when that is the main motivation behind an intervention's design. For example, there are often strong and lasting effects in terms of educational outcomes, with differences that continue

[23]For those programs that ended prior to age 5, we record achievement scores only as a long-run effect since they are typically first measured five or more years after the intervention has ended.



SOURCE: See Table 2.2.

NOTE: Number in box refers to age of child when measure was last taken. When results were mixed (gray squares), the age refers to the last age when the effect was significant. See text for full program names.

Figure 2.1—Effects of Selected Early Intervention Programs on Participating Children

throughout the school years. This finding is replicated in a number of other smaller-scale studies, as well as in pooled analyses of experimental programs (see, for example, Lazar and Darlington, 1982).

Third, there are even more white squares than filled squares. This means that most studies did not attempt to measure most of the outcomes, or that the follow-up has not been long enough to observe outcomes in many domains. As we will see later, this has implications for the cost-savings analysis: Savings for most programs will be understated because many potential benefits—especially those that can be readily monetized—were not, or have yet to be, measured.

Fourth, the benefits from early intervention programs have been measured for a diverse array of programs. Benefits are found for model programs as well as for interventions implemented on a larger scale. Significant findings are associated with programs that provide only home visits in the first few years of life, with those that offer only preschool programs for one or two years before school entry, as well as with all of the hybrid programs in between. When drawing conclusions about program benefits, it is important to keep in mind the considerable variation in the nature of the programs summarized in Figure 2.1. The programs differ from each other in such salient features as the age at program entry, duration of the intervention, the nature and quality of the services provided, and the mechanism for targeting at-risk children. With a few exceptions, it is not possible to sort out the contribution of any one of these particular design features to the strength or weakness of a particular study's findings.

These four observations, however, do not consider the variation across the domains featured in the tables. We conclude by highlighting those differences.

Cognitive and emotional development. As our review of these nine studies has demonstrated, the favorable effects in the IQ category must be interpreted with caution. Despite the recent attention on the critical period of brain development during the early years, early intervention programs—even those that provide services in the first few years of life—have been limited in their ability to demonstrate persistent cognitive effects, and the same is true for the measures of child behavior. While all of the studies in Figure 2.1 that measured IQ found significant differences during the intervention period or at

the end of the treatment, these gains disappear within a few years after the end of the program services. This conclusion is consistent with the results from other programs that we have not focused on here.[24] One of the few exceptions, the Abecedarian program—which showed significant IQ effects for children measured at age 12, four years after the end of the intervention—offers some promise that very-high-quality center-based care in the early years can produce longer-lasting cognitive benefits.

Educational outcomes. Although the IQ effects produced by early intervention programs may be short lived, there appear to be strong and longer-lasting benefits in terms of educational outcomes, such as academic achievement and other aspects of school performance. While some early intervention models, such as the IHDP and Early Training Project, found achievement effects that disappeared by the time the children reached ages 8 and 10, other programs, such as the Perry Preschool, the Chicago CPC, and the Carolina Abecedarian programs, show gains at least through age 14 or 15 (although the gap in achievement scores also tends to narrow with time).[25] When rates of special education and grade repetition or retention are reported, outcomes are always better for the treatment group, although the gap is not always statistically significant. High school graduation rates are also higher for program participants, but children have been followed long enough to observe school completion in only two cases.[26] The results from the Chicago CPC and Abecedarian eval-

---

[24]See for example, the results from the Consortium for Longitudinal Studies (Lazar and Darlington, 1982) and the other model programs and large-scale interventions reviewed by Barnett (1995) and Reynolds et al. (1997). The Milwaukee Project, which provided high-quality, full-day educational child care through age 5, is one of the only other model programs to show longer-lasting IQ effects (through age 14). It was excluded from our review because of small sample sizes (N = 20 in treatment and control groups).

[25]Of the eleven model programs with achievement data reviewed by Barnett (1995), five measured achievement test gains beyond the third grade. Two of the five programs are the Abecedarian and Perry Preschool programs reviewed here. The other three programs are the Florida Education Project, the Harlem Training Project, and the Verbal Interaction Project.

[26]Again, these conclusions are similar to those reached by Barnett (1995) based on a review of 15 model early intervention programs and 21 large-scale public early intervention programs. In a "review of reviews," Reynolds et al. (1997) also conclude that "there is substantial support for longer-term effects on children's development,

uations suggest that the gains in academic success measures may be stronger the longer the duration of the intervention program and when preschool and postschool intervention components are combined.

Criminal activity. Only four of the nine programs summarized in Figure 2.1 attempted to measure crime and delinquency behavior among youth when they were followed at older ages. In the four cases where assessments have been made, the results are generally favorable, with lower incidence and seriousness associated with juvenile offenses of those in treatment versus control groups.[27] These four programs vary in the mode, nature, and timing of service delivery, although the stronger results appear to be associated with programs that provide both high-quality day-care or preschool programs and family support services (Yoshikawa, 1995).

Employment, income, and welfare participation. Follow-up for most study participants has not been long enough to observe key economic benefits in terms of labor force activity and reliance on the social safety net as adults. The Perry Preschool program is the one exception among the studies we list in Figure 2.1, and the results are very promising. Nevertheless, until other studies are able to follow treatment and control children into young adulthood, we must attach considerable uncertainty to the ability of early intervention programs to produce differences in economic outcomes at older ages.

Health outcomes. Given the explicit focus in many programs on providing health-related services to children, it is surprising that so few programs seek to measure health benefits for children in either the short or long run.[28] The lack of long-term follow-up also pre-

---

especially for school competence (e.g., children are less likely to be retained in grade and placed in special education)" (p. 7).

[27]A recent review by Yoshikawa (1995) of over 40 early intervention programs found only 4 that explicitly measure juvenile delinquency outcomes. At the time of his review, results for the age-15 follow-up were not yet available from the Elmira PEIP. In addition to the three others we list in Figure 2.1, the fourth program identified with positive benefits in this domain was the Yale Child Welfare Project, an intervention excluded from our review because of its small sample size (N = 18 in treatment and control groups).

[28]The IHDP assessed general health status, an outcome that is not specifically classified in Figure 2.1.

cludes examining other life-course health outcomes, such as childbearing, an effect that is measured but not significant in either the Early Training Project or Perry Preschool program.

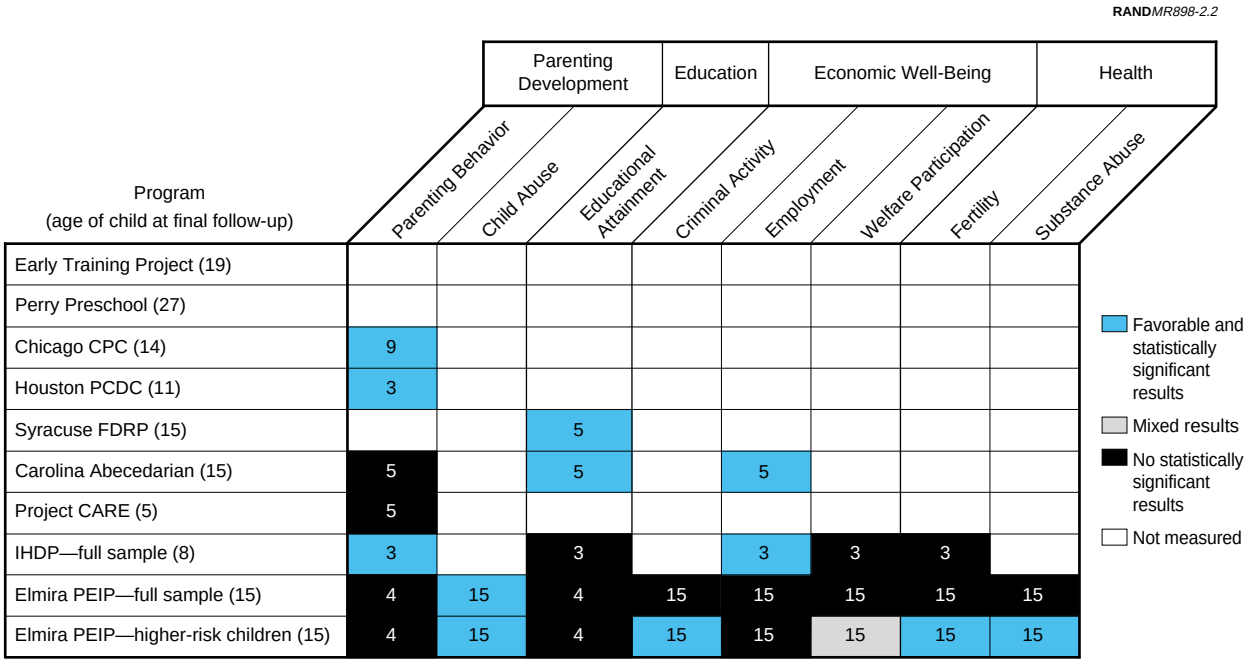
In sum, the effects of early intervention programs on children's outcomes appear to be strong in the domains of cognitive development and educational outcomes. While cognitive gains are substantial during or at the end of the intervention, they are often short-lived. However, even though IQ effects fade, other academic milestones show a greater tendency for sustained benefits. The focus on cognitive and academic measures is consistent with the motivation and design of the programs we reviewed. Other outcomes, such as economic well-being and child health, have not been assessed to the same extent—either because the follow-up period has been too short or because other factors, such as design considerations and resource constraints, led to a concentration on other outcomes.

Outcomes for Mothers

Figure 2.2 summarizes program effects on mothers. Here, in contrast to the results for children, significant favorable effects are balanced by those that are insignificant or mixed. But unmeasured outcomes dominate this figure even more than they did in Figure 2.1. While seven of the nine studies measure parental outcome in the domains covered by Figure 2.2, only four of the nine studies measure outcomes beyond parenting behavior.

Parenting. Comparisons across studies in this domain are difficult, given different measures for assessing the nature and quality of parenting. Among the studies reported in Figure 2.2, the results are mixed regarding the extent to which parenting is improved as measured by the quality of the home environment. Only one study each found benefits measured by parent-child interactions, parental involvement in schooling, or reports of child abuse and neglect. Other studies show similar results.[29] In many cases, regardless of the suc-

[29]Benasich, Brooks-Gunn, and Clewell (1982) found that 5 studies of the 27 they reviewed measured the quality of the home environment; only 2 studies found significant effects (the Houston PCDC and one other). Results for mother-child interactions were more favorable, with 10 of 11 programs that measured this outcome finding beneficial effects in program participants' favor.



SOURCE: See Table 2.3.

NOTE: Number in box refers to age of child when measure was last taken. When results were mixed (gray squares), the age refers to the last age when the effect was significant. See text for full program names.

Figure 2.2—Effects of Selected Early Intervention Programs on Participating Mothers

cess of the intervention, the home environment and quality of parenting will exert a critical influence on children's outcomes. In the Carolina Abecedarian program, for example, there was little reason to expect that the quality of the home environment for learning would be different among treatment and controls (and indeed it was not as of age 5). Nevertheless, the home environment predicted cognitive test scores for both treatment and control groups (Martin, Ramey, and Ramey, 1990). Thus, for the Abecedarian sample, the home environment remained an important influence on children's cognitive outcomes in addition to the role of the day-care intervention.

Education and economic well-being. Even fewer studies focused on maternal education and economic outcomes. Among the four programs that measure the mother's educational achievement, two find a favorable effect when the child was 5: the Syracuse home visits and day-care program and the Abecedarian day-care program. The results for employment and welfare utilization are also mixed. Mothers in the IHDP had a significantly higher level of employment (one month on average) during the three years of the intervention, while mothers in the Abecedarian program had better employment outcomes (i.e., occupational status) when the child was almost age 5. The Elmira PEIP mothers also had more months of employment over 15 years (about 7 months for the full sample and 16 months for the higher-risk sample) but the differences were not statistically significant. While the IHDP mothers showed no differences in months on public assistance during three years, mothers in the Elmira PEIP had significantly lower levels of utilization for AFDC and Food Stamps but not Medicaid. To a large extent, these results mirror those in the early intervention literature, as well as in the job training literature.[30] While evaluations have found that job training programs are largely ineffective for other segments of the population, they indicate that disadvantaged adult women benefit in terms of higher annual earnings and lower welfare participation (U.S. Department of Labor,

[30]A recent review of 27 home- and center-based programs providing early intervention services during the first 3 years of life found favorable effects on maternal employment in 10 of the 11 studies that measured the outcome, although the magnitude of the effects were modest (Benasich, Brooks-Gunn, and Clewell, 1992). Fewer studies examined public assistance utilization.

1995); for some programs, the reduction in welfare payments outweighs the cost of the training programs (Gueron, Pauly, and Lougy, 1991).[31]

Health. Among the studies summarized in Figure 2.2, the favorable health findings for mothers are limited to the Elmira PEIP, and there the benefits accrue only to higher-risk mothers. Most notable are the lower pregnancy and birth rates 15 years after the birth of their first child and the longer birth spacing between pregnancies. Given the explicit focus of the nurse home visits on maternal prenatal health behavior and postnatal life-course decisionmaking, these findings may not be replicated in other interventions that emphasize other services or alternative modes of service delivery.[32]

Given these findings, it is unfortunate that our understanding of the effects of early intervention programs on parental outcomes is so limited. Even when the child's parents are not the focus of the intervention, there may still be spillover benefits in terms of the parents' behavior. For example, programs that provide high-quality child care may increase the workforce participation of the mothers compared with mothers whose children do not participate in the program. As we will see in the next chapter, these improvements in parents' outcomes, while important in their own right, may also be important for establishing whether programs generate savings that exceed their costs.

---

[31]For reviews of the job training literature, see Reville and Klerman (1996); U.S. Department of Labor (1995), LaLonde (1995), and Grubb (1995).

[32]Benasich, Brooks-Gunn, and Clewell (1992) cite four programs in addition to the Elmira PEIP (two center-based programs and two home-based programs) that produced beneficial effects in terms of lower birth rates for participating mothers. They also find four out of a total of five programs produced beneficial effects for participants on measures of maternal mental health and self-esteem. The review of home-visiting programs by Olds and Kitzman (1993) suggests that effects on maternal life course and other maternal and child outcomes are likely to be stronger for home-visit programs that provide more comprehensive services by well-trained nurses or other professionals.