References and Notes

1. D. J. Kupfer, *Biol. Psychiatry* **11**, 159 (1976); G. Vogel, F. Vogel, R. S. McAbee, J. A. Thurmond, *Arch. Gen. Psychiatry*, in press; J. C. Gillin, W. C. Duncan, K. D. Pettigrew, B. L. Frankel, F. Snyder, *ibid.* **36**, 35 (1979).
2. N. Sitaram, R. J. Wyatt, S. Dawson, J. C. Gillin, *Science* **191**, 1281 (1976); J. C. Gillin, N. Sitaram, W. B. Mendelson, R. J. Wyatt, *Psychopharmacologia* **58**, 111 (1978); N. Sitaram, A. M. Moore, J. C. Gillin, *Nature (London)* **274**, 490 (1978).
3. J. C. Gillin, N. Sitaram, W. C. Duncan, *Psychiatr. Res.* **1**, 17 (1979).
4. D. C. Janowsky, M. K. El-Yousef, J. M. Davis, *Lancet* **1972-II**, 632 (1972).
5. N. Sitaram, A. M. Moore, J. C. Gillin, *Psychiatr. Res.* **1**, 9 (1979).
6. C. Mazure and E. S. Gershon, *Arch. Gen. Psychiatry* **36**, 521 (1979); modified from R. L. Spitzer, J. Endicott, E. Robins, *Research Diagnostic Criteria (RDC) for a Selected Group of Functional Disorders* (Biometrics Research, New York, 1975).
7. All the patients in group 1 were well known to two of the authors (J.I.N. and E.S.G.), who had followed them for more than 6 months in an outpatient lithium maintenance clinic. All patients were functioning adequately, interpersonally and in occupations commensurate with their educational background or as homemakers. During the study one patient became mildly depressed. Among the eight patients in group 2, two could be described as mildly depressed, although they were functioning adequately.
8. A. Rechtschaffen and A. Kales, *A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects* (Brain Information Service/Brain Research Institute, Los Angeles, 1968).
9. All subjects were adapted for two nights prior to the infusion studies, except for three patients (group 1) and four normal volunteers on whom we could obtain only one adaptation night. The data from these subjects were, however, no different from data from those who had been adapted for two nights.
10. T. Sagales, S. Erill, E. F. Domino, *Clin. Pharmacol. Ther.* **10**, 522 (1969).
11. Since the first non-REM period of depressed patients is short, we decided to administer our infusions during the second non-REM period. The REM-to-REM intervals and the non-REM periods subsequent to the first have previously been reported to be normal in depressed patients (*1*). To further confirm this result, we compared the second non-REM period of a group of normal subjects used in previous experiments (*2*) (N = 21, mean = 84.7 minutes) with that of 47 patients with primary depressive illness (mean = 81.9 minutes) who had been studied in our laboratory; there was no significant difference.
12. N. Sitaram, R. L. Cohen, A. Rey, D. L. Murphy, J. C. Gillin, unpublished data. Seven depressed patients (five female and two male, age = 26.8 ± 6.1 years) were tested with the 0.5-mg arecoline dose. The inf-REM_2 latency for the depressed patients (10.5 ± 11.1 minutes) was significantly lower than in the normals (37.5 ± 21.9 minutes) ($P < .02$, two-tailed Mann-Whitney U test). All subjects had been drug free for at least 2 weeks before testing.
13. M. Jouvet, in *Cholinergic Mechanisms*, P. G. Waser, Ed. (Raven, New York, 1975), pp. 455–476.
14. J. C. Gillin, W. B. Mendelson, N. Sitaram, R. J. Wyatt, *Ann. Rev. Pharm. Toxicol.* **18**, 563 (1978); A. G. Karczmar, *Brain Acetylcholine and Neuropsychiatric Disease*, K. L. Davis and P. A. Berger, Eds. (Plenum, New York, 1978) pp. 265–310; J. A. Hobson, R. W. McCarley, T. M. McKenna, *Prog. Neurobiol. (N.Y.)* **6**, 155 (1976); G. W. Vogel, *Am. J. Psychiatry* **135**, 1531 (1978).
15. K. L. Davis, P. A. Berger, L. E. Hollister, E. DeFraites, *Arch. Gen. Psychiatry* **35**, 119 (1978); D. S. Janowsky, M. K. El-Jousef, J. M. Davis, H. J. Sekerke, *ibid.* **28**, 542 (1973); K. L. Davis, L. E. Hollister, J. Overall, A. Johnson, K. Train, *Psychopharamcology* **51**, 23 (1976); C. Tamminga, R. C. Smith, S. Chang, J. S. Haraszti, J. M. Davis, *Lancet* **1976-II**, 905 (1976).

17 October 1979; revised 2 January 1980

Preschool Programs and Later School Competence of Children from Low-Income Families

Abstract. *At follow-up in 1976, low-income children who had attended infant and preschool programs in the 1960's had significantly higher rates of meeting school requirements than did controls, as measured by lower frequency of placement in special education classes and of being retained in grade (held back).*

Controversy over the effectiveness of preschool programs for low-income children has been extensive. This report evaluates the effectiveness of preschool on the basis of a new specification of a commonly accepted goal, improving later school competence.

Preschool programs for low-income children developed rapidly in the 1960's. Project Head Start, initiated in 1965, signaled a nationwide commitment to such programs. It was designed to improve children's intellectual skills, to foster their emotional and social development, to help meet their health and nutritional needs, and to involve parents and the larger community in those purposes (*1*). The high educational expectations were dampened by evaluations, such as the Westinghouse-Ohio University report, which concluded that preschool did not have a lasting effect on children's intellectual development (*2, 3*). Among social scientists the negative results occasioned a review of the premises of early intervention, a reopening of the nature-nurture controversy, inquiries into the nature and meaning of intelligence testing, and doubts about the role of early experience in later development (*4*).

The research reported here provides evidence of the effectiveness of preschool programs for children from low-income families. [Longer reports are available (*5–7*).] It is based on follow-up of eight separate preschool research projects, and it examines their effects well beyond the primary grades: the median subjects were in the seventh grade.

The research was conducted by members of the Consortium for Longitudinal Studies. Twelve investigators (*8*) who had independently designed and conducted experimental preschool programs in the 1960's pooled their original data and conducted a common follow-up study in 1976–77 of the original participants and control groups. Two consortium members (*9*) who had not conducted preschool programs coordinated the follow-up work.

The original preschool programs were located at 11 urban and rural sites in the Northeast, Southeast, and Midwest. Their delivery systems and curricula varied considerably. Some programs concentrated on home visits to teach mothers how to be more effective teachers of their infants and toddlers; in others, 3- or 4-year-olds were taught in a group setting; others had both center and home-visit components. Curricula in the centers were based on the Bank Street child-development model, on Montessori's methods, on Piagetian theory, on the Bereiter-Engelmann method, and others.

The preschool programs of the consortium investigators were originally designed for both demonstration and research purposes. A subset was designed specifically to answer the question of whether preschool was effective, with random assignment (or a close approximation) of children to treatment (preschool) and control groups. We have designated these as experimental designs. Other programs were designed primarily to answer questions such as which curricula were most effective and what were optimal ages of entry or length of stay in a program. Matched comparison groups were used in those designs; they are here designated quasi-experimental.

Of approximately 2700 subjects and controls in the 11 original projects, 1599, ranging in age from 9 through 19 years, were available on follow-up. They can be generally characterized as blacks (94 percent) from low-income families in which the mothers had completed 10.3 years of education on the average and the average head of household was a semiskilled or unskilled worker. Other characteristics are summarized in Table 1. Five types of follow-up data were collected in a common format: an individually administered intelligence test (Wechsler), school record information, scores on school-administered standardized achievement tests, and interviews with participants and parents (usually the mothers).

To guard against artifactual results due to differences between the programs, the subjects were never pooled into a single sample. In all hypothesis tests treatment children were compared with control children from the same project site; the

0036-8075/80/0411-0202 $00.50/0

results were pooled by means of the pooled z technique (*10*). This method tests the null hypothesis that on the average preschool has no effect on later school performance (*11*). To ensure that no single program accounted for a significant pooled result, we deleted the program with the strongest result from each analysis and recalculated the pooled z score. Projects with designs most appropriate to test the effectiveness of preschool were analyzed separately from the quasi-experimental projects. The P values reported here are two-tailed.

Attrition due to mobility, death, or other events is a serious threat to the validity of a longitudinal study. The median recovery rate for the seven projects used in the pooled z analyses reported here was 74 percent (*12*). Virtually all attrition was due to inability to find subjects; less than 3 percent of the recovered subjects refused to participate.

Analyses were performed to assess whether the total samples found on follow-up differed from the total original samples, whether the found treatment group subjects differed for the found control group subjects, and whether the pattern of attrition differed for treatment and control groups. Attrition appeared to be essentially random, introducing no known biases into our results.

Our findings replicated results on IQ reported by numerous previous investigators—large effects in tests given soon after the program, tapering down to smaller but statistically significant effects 3 or 4 years after preschool and vanishing thereafter. This finding has a positive aspect: if 1 year of enrichment still has effects 3 or 4 years later, then 12 years of enriched schooling might have lifelong effects. IQ's, however, are merely indirect predictors of school success; our most important analyses concerned the effect of preschool on direct measures of school success.

From school records field workers reported whether a child had been retained in grade (held back) or assigned to special education classes at any time during his or her school career. The category of special education included classrooms for children diagnosed as educable or trainable mentally retarded, learning-disabled, or emotionally disturbed. Classes for children with speech and hearing difficulties were not considered special education programs for these analyses.

Local school districts have varying policies regarding special education classes and grade retention; however, the two indicators share the implication that the student has not met the school's requirements for normal progression. Therefore, we combined the two indicators into a single measure distinguishing students who met from those who failed to meet the school's basic requirements; we counted as meeting those requirements children who were never placed in special education classes or retained in grade.

From seven projects with appropriate data (*13*), children who attended preschool were found to be more likely to meet the requirements of their schools than children with no preschool experience (Table 2). The median rate of failure by that standard was 45 percent in the control group, 24 percent in the treatment group. The difference was statistically significant for the seven projects (pooled $z = 3.59$, $P < .001$) and robust ($P = .006$ after the strongest result is deleted from the pooled z). The result was even stronger for the projects with experimental research designs (pooled $z = 4.00$).

Preschool apparently had a stronger effect on rate of assignment to special education than it had on grade retention. For the six projects in which special education data were collected, the median rate of assignment to special classes at any time in the school career was 29 percent in the control group, 14 percent in the treatment group. The preschool children had statistically significant lower rates of assignment to special classes in four of the six projects (*14*). The combined result for the six projects was statistically highly significant (pooled

Table 1. Characteristics of the follow-up samples, in percent.

Project	N	Attended preschool	Male	Black	Two-parent families	Mean education of mother*	Mean preprogram IQ†
Beller	119	36	49	92	72	10.5	91.1
Deutsch	54	74	44	100	59	9.9	93.9
Gordon	107	78	47	94	57	9.9	
Gray	75	48	44	100	71	9.2	88.2
Karnes	151	100	50	66	61	10.1	92.1
Levenstein	154	68	55	94	62	10.6	84.3
Miller	139	77	46	93	46	10.7	93.1
Palmer	224	79	100	100	73	11.1	92.2
Weikart	123	47	59	100	53	9.4	79.0
Woolman	297	66	57				71.5
Zigler	156	54	56	89	77	11.3	
Median	1599‡	68	50	94	61.5	10.3	92.1

*Grades completed. †Stanford-Binet except in Levenstein (Peabody Picture Vocabulary) and Woolman (WISC). The median does not include Levenstein and Woolman scores. ‡Total N school record data.

Table 2. Comparison of treatment and control groups on failures to meet school requirements (as evidenced by being placed in special education classes or being retained in grade). Most of the children in the respective projects were in the following grades when the data were collected: Gordon, 5th; Gray, 12th; Palmer, 7th; Weikart, 4th; Beller, 12th; Levenstein, 3rd; Miller, 7th. Certain control groups were excluded so that the comparisons would be as close to experimental as possible.

Project	N	Failed to meet requirements (%)		Chi-square	P (two-tail)	Pooled z	Pooled P (two-tail)
		Treatment	Control				
			Experimental design				
Gordon	82	39.1	61.5	2.25	.134		
Gray	55	52.8	68.4	1.25	.263		
Palmer	221	24.1	44.7	7.66	.006		
Weikart	123	17.2	38.5	6.78	.009		
Median	481	31.6	53.1			4.00	< .001
			Quasi-experimental design				
Beller	69	45.9	50.0	.11	.737		
Levenstein	127	22.1	43.5	4.47	.035		
Miller	125	20.6	11.1	.89	.346		
			All projects				
Median	802	24.1	44.7			3.59	< .001

$z = 3.52$, $P < .001$) and was robust ($P = .010$ after deleting the strongest result). For the three programs with experimental research designs the combined result was even stronger (pooled $z = 4.04$, pooled $P < .001$).

An analysis of the special education data by program variation confirms this result. All but one of the projects with special education data included more than one variation of the preschool program. There were altogether seven different program variations which could be compared with well-matched control groups on frequency of placement in regular classes instead of special education classes. In all seven programs the proportion of program children who were always placed in regular classes exceeded the proportion of control group children so placed. The probability that this would occur by chance is .020 (*5*). The seven programs were ranked on effectiveness as measured by the difference between those proportions. In the median program on this ranking, 29.4 percent of the control group children but only 5.3 percent of the experimental group children were placed in special education classes.

Across the eight projects with appropriate data, treatment children were less likely to experience grade retention than were controls (pooled $z = 2.36$, pooled $P = .018$), with stronger results for projects with experimental research designs (pooled $z = 2.87$, pooled $P = .004$) (*15*). Neither result was robust, however. Across the eight projects the median rate of grade retention was 25 percent in the treatment group and 31 percent in the control group; for the four experimental projects it was 26 percent in the treatment group and 37 percent in the control group.

Treatment and control groups generally did not differ significantly on background or demographic variables measured before preschool, but standard multiple regression was used to control for whatever differences were observed. Covariates controlled in this way were level of maternal education (grades completed), presence versus absence of father, number of siblings, sex, ethnicity (black versus other), and preprogram IQ. On the average, with these initial ability and home background factors controlled, 16 percent more of the students with preschool experience than those without were always on grade and in regular classes. Results were significant and robust for all seven projects and strongest for the group of experimental projects (pooled $P < .001$; .016 with robustness test).

We also considered whether certain kinds of children benefited from preschool—for example, those with higher initial ability or those from families of somewhat higher socioeconomic status (*3, 16*). The interaction of treatment/control status with sex, ethnic background, preprogram IQ, presence versus absence of father, number of siblings, and level of maternal education (all measured prior to preschool) was examined in a regression format. No significant interaction effects were found.

RICHARD B. DARLINGTON
Department of Psychology,
Cornell University,
Ithaca, New York 14853

JACQUELINE M. ROYCE
ANN STANTON SNIPPER
HARRY W. MURRAY
IRVING LAZAR
Department of Human Service Studies,
Cornell University

References and Notes

1. L-E. Datta, in *Studies in Evaluation Research*, C. Abt, Ed. (Sage, Beverly Hills, Calif., 1976). For theoretical work emphasizing the crucial nature of early experience for the course of later development, see J. McV. Hunt, *Intelligence and Experience* (Ronald, New York, 1961); B. S. Bloom, *Stability and Change in Human Characteristics* (Wiley, New York, 1964); D. O. Hebb, *The Organization of Behavior* (Wiley, New York, 1949).
2. Westinghouse Learning Corporation/Ohio University, *The Impact of Head Start: An Evaluation of Head Start on Children's Cognitive and Affective Development* (National Technical Information Service, Springfield, Va., 1969); F. D. Horowitz and L. Y. Paden, in *Child Development and Social Policy*, B. M. Caldwell and H. N. Ricciuti, Eds. (Univ. of Chicago Press, Chicago, 1973).
3. U. Bronfenbrenner, *DHEW Publ. No. (OHD) 74-25* (1974).
4. A. Jensen, *Harv. Educ. Rev.* **39**, 1 (1969); M. Cole and J. Bruner, *Am. Psychol.* **25**, 867 (1971); A. M. Clarke and A. D. B. Clarke, *Early Experience: Myth and Evidence* (Free Press, New York, 1976).
5. I. Lazar, J. Royce, H. Murray, A. Snipper, R. Darlington, *Monogr. Soc. Res. Child Dev.*, in press.
6. Consortium on Developmental Continuity, *DHEW Publ. No. (OHDS) 78-30130* (1977).
7. Consortium for Longitudinal Studies, *DHEW Publ. No. (OHDS) 79-30178* (1978).
8. E. K. Beller, M. Deutsch, C. Deutsch, S. Gray, I. Gordon (deceased 1978, replaced by R. Jester), M. Karnes, P. Levenstein, L. Miller, F. Palmer, D. Weikart, M. Woolman, and E. Zigler. For details on individual projects see chapters by I. Gordon and R. Jester; M. Karnes, R. Zehrbach, J. Teska; P. Levenstein; F. Palmer and R. Siegel, in *The Preschool in Action: Exploring Early Childhood Programs*, M. Day and R. Parker, Eds. (Allyn & Bacon, ed. 2, Boston, 1977); E. K. Beller, in *Handbook of Research on Teaching*, R. Travers, Ed. (Rand-McNally, Chicago, 1973); M. Deutsch, E. Taleporos, J. Victor, in *DHEW Publ. No. (OHD) 74-24*, S. Ryan, Ed. (1974); S. Gray and R. Klaus, *Child Dev.* **41** (1970); L. Miller and J. Dyer, *Monogr. Soc. Res. Child Dev.* (1975), vol. 40; D. Weikart, J. Bond, J. McNeil, *Monographs of the High/Scope Educational Research Foundation* (1978), vol. 3. M. Woolman, *Learning for Cognition* (report to the New Jersey Department of Education, 1971).
9. I. Lazar and R. B. Darlington.
10. F. Mosteller and R. R. Bush, in *Handbook of Social Psychology*, G. Lindzey, Ed. (Addison-Wesley, Reading, Mass., 1954). For each project the exact P value of the result was calculated and converted to a z score with the sign corresponding to the direction of the effect. The z scores were then summed: $z = \Sigma z_i/\sqrt{k}$, where z_i is the z score from project i and k is the number of projects. The significance of the pooled z score was then determined. Since contrary results in different projects cancel one another out, this tests the "average" effect of preschool. When results were tested by chi-square, uncorrected chi-squares were calculated [G. Camilli and K. Hopkins, *Psychol. Bull.* **85**, 1 (1978)].
11. A less conservative null hypothesis would be: preschool has no effect on children's later school performance. This hypothesis could be rejected on the basis of the successful outcome of a single program.
12. Rates of recovery were: Beller, 66 percent; Gordon, 35 percent; Gray, 85 percent; Levenstein, 75 percent; Miller, 51 percent; Palmer, 74 percent; and Weikart, 100 percent. Gordon's recovery efforts were hampered by budget.
13. The various analyses included different numbers of projects because of the special nature of Zigler and Palmer data. Zigler collected data on retention but not special education placements. Palmer did not distinguish between special education placements and grade retentions; nevertheless, his data were included in the retention analysis as well as the composite analysis because he reported that almost all failure to meet school requirements were retentions. Three additional projects collected school data but were not included in any of these analyses. The Deutsch project encountered exceptional problems of attrition because of New York City school reorganization; the data were analyzed but not pooled with other projects (*7*). The Woolman project treated all high-risk children in their school district; the control group consisted of a random sample of the children who had entered school in the previous year. Karnes compared five different preschool curricula and did not have a control group. For Woolman and Karnes results see (*6*).
14. P values for individual projects were: Gordon, .024; Gray, .004; Weikart, .060; and Levenstein, .005. Results for Miller ($P = .466$) and Beller ($P = .632$) favored controls. Miller's control group was discovered, post hoc, to contain a significantly higher percentage of white children and two-parent families than the treatment group did, and income of control families was somewhat higher than that of treatment families. The schools attended by Beller subjects apparently rarely placed children in special classes (only one control child and two treatment children had been assigned to special classes in the 12 years of schooling).
15. Children who had been in special education classes and had never been retained in grade were excluded from the retention analysis. P values for individual projects were: Gordon, .956; Gray, .291; Palmer, .006; Weikart, .065; Beller, .475; Levenstein, .531; Miller, .248 (controls had lower grade-retention rate); and Zigler, .452.
16. E. Herzog, C. H. Newcomb, I. H. Cisin, in *DHEW Publ. No. (OHD) 74-24*, S. Ryan, Ed. (1974); H. Beilin, in *Preschool Programs for the Disadvantaged*, J. C. Stanley, Ed. (Johns Hopkins Press, Baltimore, 1972).
17. We thank members of the Consortium for Longitudinal Studies for their indispensable cooperation; E. Grotberg and B. Brown, of the Administration for Children, Youth and Families, for their encouragement; and our colleagues at individual project sites and R. Hubbell, D. Koretz, J. Vopava, M. Rosché, and R. Snipper for important contributions to this work. Supported by OHDS, DHEW grants Nos. 90-C-1311 and 18-76-07843, Maternal and Child Health grant No. MCT-004012-01-0, and by the William and Flora Hewlett Foundation.

2 November 1979; revised 4 February 1980