

Baumeister, A.A., & Bacharach, V.R. (1996). A critical analysis of the infant health and development program. *Intelligence*, 23(2): 79-103. (Sep-Oct 1996) Published by Elsevier Science Limited (ISSN: 1873-7935).

A Critical Analysis of the Infant Health and Development Program

Alfred A. Baumeister and Verne R. Bacharach

ABSTRACT

The largest study designed to ameliorate adverse effects of premature low birth weight (LBW) and to prevent mental retardation is the Infant Health and Development Program (IHDP). This was a randomized, multisite intervention: home visits for 3 years, parents meetings, and intensive preschool education for 2 years. IHDP reported results alleging the program significantly influenced intelligence and prevented mental retardation. We conducted an independent analysis of the original computerized database (at 3 years). Five-year follow-up data were obtained from the journal publication and from data on file with the National Auxiliary Publication Service. Our intent was to determine the magnitude, durability, and clinical significance of purported intervention effects and how these are mediated. Methods used were primarily multivariate correlational analyses and examination of the logic underlying the conclusions. Results suggest alternative interpretations of claims regarding IHDP. Effects are explained by confounding variables, questionable analytical procedures, distorted interpretations, and data inconsistencies. Effect sizes and specificity of effect reported by IHDP do not survive scrutiny either in the original database or at 5 years. Given the vastly complex nature of premature LBW, IHDP was poorly conceived, failing to produce meaningful and enduring effects on IQ. Policy conclusions for interventions with LBW infants stemming from IHDP are misleading.

Most low-birth-weight children (LBW, < 2500 g) appear to develop normally. However, the literature is replete with studies showing that on the whole LBW infants, particularly those born prematurely and extremely small, are at substantially elevated risk for mortality, and surviving, for myriad health, neurological, developmental, behavioral, social, learning, and other problems. As a group, LBW infants present with neurodevelopmental handicaps at three times the rate of normal-birth-weight (NBW) babies. These include chronic disabilities such as cerebral palsy, autism, mental retardation, developmental delays, hearing and visual impairments, and mental disorders. LBW children between ages 6 to 15 years are 50% more likely than NBW children to be enrolled in a special education program even after controlling for regional, family, and individual factors (Chaikind & Corman, 1991).

Not surprisingly, a weight-risk gradient defines the relationship in that the lower the birth weight within the LBW range (< 2500 g) the greater the risks in terms of likelihood of occurrence, severity of outcome, and number and kind of adverse health and functional disabilities (Alberman, 1994; Escobar, Littenberg, & Petitti, 1991; Hack, Klein, & Taylor, 1995; McCormick, 1989; McCormick, Brooks-Gunn, Workman-Daniels, Turner & Peckham, 1992). Across a broad LBW span there is a linear relationship between birth weight and IQ for children at 6 years (Breslau et al., 1994). Studies of very low (< 1500 g) and extremely LBW (< 1000 g) survivors indicate that at least 25% show significant developmental handicaps (Escobar, Littenberg, & Petitti, 1991), whereas 40% to 65% present learning disabilities, hyperactivity, attention deficit disorder, and other problems (Baumeister, Kupstas, & Woodley-Zanthos, 1993). Of the 75% who appear to have developed intact during infancy and whose neurological or developmental impairments are not so obvious, about half will still require special services (Klein, Hack, & Breslau, 1989).

The high relative long-term risk for health and behavioral problems at school age among extreme LBW babies (< 1000 g) is confirmed by a number of studies, from four countries, in which children were followed beyond 5 years (e.g., Hille et al., 1994; Klebanov, Brooks-Gunn, & McCormick, 1994; Saigal, Szatmari, Rosenbaum, Campbell, & King, 1991; Victorian Infant Collaborative Study Group, 1991).

Developmental risks attendant to LBW are strongly associated with demographic factors such as poverty, social class, race, marital status, maternal age, and maternal education. But even with such variables controlled, statistically or by sampling, there remains a direct and major risk for poor developmental outcome among LBW infants, especially those born weighing 1500 g or less.

PUBLIC HEALTH ASPECTS

From a public health perspective, the problem of LBW takes on special significance in that (a) LBW rates have been gradually increasing in the United States since reaching a low in 1984 of 6.7% to 7.3% in 1994 (Ventura, Martin, Mathews, & Clarke, 1996); (b) there are huge disparities of relative risk of LBW associated with race, maternal age, parental education, and socioeconomic status (Baumeister,

Kupstas, & Woodley-Zanthos, 1993); (c) the United States ranks significantly behind (14 in 1980, 17 in 1989) many other countries in incidence of LBW per 1000 births, especially very LBW (Baumeister & Kupstas, 1990; Miller, Fine, & Adams-Taylor, 1989); (d) with application of new technologies in neonatology, especially treatment with exogenous surfactant (and, more recently, birthing mothers with corticosteroids), there has been a major decline in pulmonary-related mortality with a concomitant increase in prevalence (but not necessarily incidence) of those with handicapping conditions (McCormick, 1993); (e) costs associated with initial hospitalization, rehospitalization, related long-term medical, social, and educational interventions are enormous, approaching \$4 billion annually (in 1988 dollars) just for health care and incremental costs during the 1st year of life (Lewit, Baker, Corman, & Shiono, 1995); and (f) given the fiscal reality of escalating costs to save seriously compromised ultralight babies in the face of a looming scarcity of resources and in consideration of quality-of-life issues, in one guise or another rationing will surely become a divisive legal and ethical dilemma (Eddy, 1994; Tyson, Perlman, Rosenfeld, & Arencibia-Mireles, 1992). (Note the furor over the much-publicized Oregon Medicaid waiver.)

Clearly, premature LBW is a major and complex public health problem that presents numerous challenges, both in epidemiological and clinical domains. The literature contains many accounts of methods designed to prevent or minimize adverse sequelae. Approaches to this end have focused on prenatal care, infant stimulation, family support, and early preschool interventions. In public health terms the fundamental question becomes one of determining which is the most promising avenue of intervention to avert or reduce deleterious effects associated with LBW.

Prenatal Care

A report by the Institute of Medicine (Institute of Medicine, 1985) concluded that, after socioeconomic status, prenatal care is the most important factor influencing birth outcome (including prematurity or LBW). Other reviewers have also concluded that prenatal care improves birth outcomes, particularly premature LBW (e.g., Klein & Goldenberg, 1993).

Even though adequate prenatal care has long been accepted on its intrinsic merits as a primary preventive measure, there are difficulties in determining precisely the benefits of antenatal care, especially in uncomplicated pregnancies (Alexander & Korenbrot, 1995; Baumeister, Kupstas, & Woodley-Zanthos, 1993; Fiscella, 1995). Obviously, medical services cannot be ethically withheld from a control group. In 1992 over 98% of pregnant women received some prenatal care (National Center for Health Statistics, 1994), although for many mothers the level of care is inadequate when measured in terms of timing (third trimester or not at all), particularly among minorities, poor, very young women, unwed mothers, and those with less than 12 years of education (National Center for Health Statistics, 1993, 1996).

Evidence in favor of effectiveness of standard prenatal care for reducing incidence of prematurity and LBW is equivocal, focusing as it does on the detection of major complications of pregnancy. Cohort studies designed to evaluate differences

in birth outcome, including LBW, comparing standard prenatal care with a more comprehensive system of supports have generally produced negative results, although retrospective studies are somewhat more positive (Baumeister, Kupstas, & Woodley-Zanthos, 1993). But given underlying risk factors for certain well-defined populations, augmented early prenatal care (including social, educational, and nutritional services) for specific subgroups of high-risk women who have preventable risks for LBW may be feasible (McLaughlin et al., 1992; alds, Henderson, Tatelbaum, & Chamberlin, 1986). Does good prenatal care (i.e., the focus by definition is on postconceptual and often immutable risk) convert into preconception care and modification of risk factors for the next pregnancy (Jack & Culpepper, 1990)? Another possibility is to devise systems approaches to delivery of quality and specialized prenatal and preconceptional care, including case management and comprehensive ancillary services, to those women who might otherwise not receive any or for whom monitoring would be substandard (Meis et al., 1987; U. S. Department of Health and Human Services, 1989).

Infant Stimulation

Another strategy is to provide early opportunities or experiences that will forestall adverse outcomes for babies born prematurely and small. Some of these are short-term interventions, usually in the neonatal intensive care unit (NICU), designed to structure stimulation (visual, auditory, tactile, or social) during the immediate post- and neonatal periods (Bennett, 1987). Others have focused on the mother to enhance her appreciation of the unique characteristics of the LBW infant and to teach her how to react appropriately to create a healthy transaction with the child (Rauh, Nurcombe, Achenbach, & Howell, 1990; Zahr, Parker, & Cole, 1992). A few studies have combined infant stimulation and parent training. A recent study of environmental neonatology has shown that highly individualized, comprehensive, and family-oriented NICU care reduced medical and neurodevelopmental sequelae (included were Bayley scores at 9 months) among babies weighing less than 1000 g (Als et al., 1994). Nevertheless, taken as a whole, this research has produced findings that are extremely mixed and usually without long-term follow-up (Wolke, 1991).

THE INFANT HEALTH AND DEVELOPMENT PROGRAM (IHDP)[1]

Still another avenue is to provide intensive early preschool interventions based on models designed for NBW infants from socially disadvantageous circumstances. The first large-scale clinical trial to assess the efficacy of this type of intervention is the Infant Health and Development Program (Gross, 1992; IHDP, 1990). Randomized intervention trials involving 985 LBW premature babies[2] were conducted at eight sites throughout the United States. Randomization procedures involved stratification and other considerations (Kraemer & Fendt, 1990). The babies were stratified into two weight groups: *heavier* (one third weighed 2001 to 2500 g) and *lighter* (two thirds were 2000 g or less). About one third of the children were assigned to an intervention group; the remainder to a follow-up group.

The program covered a 3-year span. Upon discharge from the nursery all children received medical, social, and developmental assessments with referral as indicated. Those in the intervention condition received a package of services including home visitation and, at 12 months, parent group meetings and centerbased education. The primary dependent measures were cognitive status, behavioral competence, and health status. A more complete description of the project can be found in IHDP (1990) and Gross (1992).

The Reported Results (1990)

Cognitive Status. Numerous articles have appeared reporting results of the IHDP on such measures as cognitive development, behavioral deviations, health indices, and maternal employment (Brooks-Gunn, Gross, Kraemer, Spiker, & Shapiro, 1992; Brooks-Gunn, Liaw, & Klebanov, 1992; Brooks-Gunn, Klebanov, Liaw, & Spiker, 1993; Brooks-Gunn, McCormick, Shapiro, Benasich, & Black, 1994; McCormick et al., 1991; McCormick, McCarton, Tonascia, & Brooks-Gunn, 1993; Ramey et al., 1992; Ramey & Ramey, 1992a, 1992b; Spiker, Ferguson, & Brooks-Gunn, 1993). The primary outcome that has received the greatest attention is, understandably, cognitive development (DQs/IQs). The first published report (i.e., after the intervention families had been enrolled for 3 years; 2 years in the day-care component) revealed significantly higher mean Stanford-Binet scores for the intervention than for the follow-up children (93.5 vs 84.5, Table 6, p. 3040). This is an effect size (*ES*) of .46 (*N* = 908).

However, there was a distinct weight gradient effect in that those at higher birth weights showed the most pronounced intervention effects. The difference between the intervention and follow-up children in the "heavier" (2001-2500 g) stratum was 13.2 IQ points, *ES* = .83, *N* = 347, on the Stanford-Binet; for the "lighter" (< 2000 g) the difference was only 6.6 points, *ES* = .41, *N* = 561. Mean IQs for babies 1500 g and below were 85.3 and 80.2 for the intervention (*N* = 82) and follow-up groups (*N* = 150), respectively. (These means were not reported in the original publication; we obtained them as part of our reanalysis of the database.) In short, the intervention was more efficacious at 36 months for higher birth-weight infants (2001 to 2500 g).

Only in the case of the heavier (2001-2500 g) infants did the effect size reach their own preset criterion: "An effect size of 0.5 ... was considered critical" (Kraemer, 1988, p. 14). Of the 11 effect sizes reported by IHDP (1990, Table 3) for the primary outcome measures, only one (36-month IQs for the heavier group) reached that criterion.

Cognitive assessments were conducted when the day program began at 12 months, at 24 months, and at 36 months when the program ended. Brooks-Gunn, Klebanov, Liaw, and Spiker (1993) presented a detailed analysis of cognitive development of children from both arms of the study in three weight strata. Both intervention and follow-up groups declined in test scores over time; the effect was greater for the follow-up infants. This was a three-way interaction in that the treatment effect was contingent on birth weight and the group disparity was greatest among children weighing more than 2000 g at birth. Brooks-Gunn et al.

caution that "follow-up is necessary to elucidate the continuing developmental trajectory of these two groups" (1993, p. 747).

Large differences in overall IQ effects were also found for the different sites: follow-up groups' mean IQs ranged from 96.7 for Boston to 68.0 for Miami. In addition, treatment effects were minimal at some sites—for example, Boston, a difference of .4, and more pronounced at others, for example, Little Rock, Arkansas, 14.3—probably reflecting differential parental education and economic status. As reported, the two-way Treatment x Site interaction was not statistically significant after controlling for other initial status variables. Significant IQ differences, in the expected directions, were found for other initial status variables: race, maternal education, and a neonatal health index. However, the only initial variable found to interact with intervention was birth weight—in other words, the more compromised the child by circumstances of LBW, the smaller the treatment effect.

When site and initial status variables were controlled, 3-year IQs of less than 70, which is in the conventional mental retardation range, were 2.7 times greater in the follow-up group. However, for infants at or less than 1500 g there was no differential relative risk for mental retardation associated with the intervention.

Other Primary Outcomes. Although the main thrust of IHDP was on cognitive development, that is, IQs, measures were obtained on "behavioral competence" (mother's report of behavioral problems on the Child Behavior Checklist) and "health status" including maternal report of morbidity, maternal report of functional limitations, growth parameters, and the General Health Ratings Index. Six health status measures were reported (IHDP, 1990).

Reported behavioral problems were 1.8 times greater, $ES = .18$, in the follow-up group (IHDP, 1990). The only variable found to interact with intervention was maternal education, $p = .009$; that is, no difference in reported behavioral problems between groups for more highly educated mothers (some college education), whereas less educated mothers in the intervention condition reported fewer behavioral problems than did less educated mothers in the follow-up condition. The only health status measure to show a significant treatment effect was mother's report on the Morbidity Index for the lighter-born children, $ES = .29$. (This was a highly questionable measure that could include a number of different events and could range only from 0 to 3.) Maternal age interacted with treatment such that younger mothers in the intervention group reported more frequent adverse health conditions than their counterparts in the follow-up cohort, $p = .003$.

Conclusions Reached. Based on their original analysis (when the children were 36 months of age) a number of conclusions having distinct programmatic, health, and policy implications were suggested by the members of the IHDP (1990). Perhaps the most overarching is that the "comprehensive and intensive early intervention program shows substantive promise of decreasing the number of LBW premature infants at risk for later developmental disability" (p. 3041). This study is seen as "timely and relevant" to policy concerns at federal and state levels for interventions designed for children who are at risk for developmental delay.

A REANALYSIS OF THE IHDP DATABASE AT THREE YEARS

The IHDP database is large, complex, available,[3] and invites independent analysis, particularly in view of the significant policy and programmatic implications identified in the original report (IHDP, 1990) and in many subsequent publications. In an editorial in the same issue of the original report of IHDP (1990) in the *Journal of the American Medical Association*, Richmond stated that "the results have considerable clinical and potential public policy significance" (Richmond, 1990, p. 3069). Such sentiments may be premature and unwarranted given a careful analysis of the original IHDP database and examination of longer-term effects (Brooks-Gunn, McCarton, et al., 1994).

Rationale

The original analysis as described by IHDP (1990) was repeated as a check to ensure that we were looking at the same numbers and could (and did) replicate the analyses as reported in IHDP (1990). However, we also approached the data from different multivariate and theoretical perspectives.

The major thrusts of our reanalysis of the IHOP database were to (1) evaluate exactly how large the intervention effect is; (2) control for important variables that mediate whatever effects are observed, and for whom; (3) determine which intervention components had a direct effect, if any, on outcome; (4) determine where in the multistage intervention program effects, if any, occurred. Certain claims made in regard to this IHOP during the first 3 years are evaluated in light of this reanalysis.

The IHOP authors claim that at 36 months children in the intervention group, as a whole, tested significantly higher on the Stanford-Binet than their follow-up counterparts (IHOP, 1990, p. 3038). But there are serious interpretative questions as to the validity of these reported effects, how they are mediated, when they began, whether they can be sustained, and what they mean medically, educationally, and socially. These are considerations that become all the more salient in light of the policy determination by the Centers for Disease Control and Prevention to initiate programs based on the IHOP model.

RESULTS

Just how powerful is the intervention, taking into account the reported birthweight interaction (IHOP, 1990)? Our analyses were conducted on children for whom, in addition to IQ scores, there were complete data available with respect to all selected variables in the particular analysis. Following are zero-order correlations of child IQ at 36 months (for the pooled sample of intervention and control groups with $N = 775$) in order of magnitude, with variables known (in this case, including the treatment) to influence child IQ: maternal IQ = .50; home environment (at 12 months) = .45; maternal race (Black vs all others) = .34; maternal education (high school or more versus less than high school) = .27; intervention = .21; and birth weight = .12.

Table 1 presents the results of this analysis including the correlations, beta, standard error of beta, and, where appropriate, *ES*. Effect sizes presented here are slightly different from those reported previously because calculations were based on 775 mothers whose IQs were coded "valid" as defined by IHOP.

TABLE 1
Zero-Order Correlations of Variables Associated With Child IQ (*N* = 775)

	<i>r</i>	<i>b</i>	<i>SE_b</i>	<i>ES</i>
Maternal IQ	.50	.47	.03	
Home (Range = 7-45)	.45	1.42	.10	
Race (Black/other)	.34	13.36	1.33	.72
Education (high school+ /no high school)	.27	11.14	1.41	.57
Intervention	.21	8.39	1.41	.42
Weight (in 100 g increments)	.12	.54	.16	

The zero-order correlation between intervention and IQ at 36 months accounts for but 4.4% of the IQ variance in this study, less than the other variables except for initial birth weight. Fifty percent of child IQ variance is associated with maternal IQ variance in the IHDP study. This is a kinship correlation so that it represents, without squaring, the percentage of variance explained by hereditary factors and shared environment (Jensen, 1971). Compare the magnitude of the intervention effects at 36 months with such factors as maternal IQ, maternal race, and the home environment. Heritability in this case is precisely what would be expected. Of course, this does not mean that changes in the environment could not affect mean group differences. In the case of IHDP, the intervention effect was simply small.

The Birth Weight by Intervention Interaction

A great deal was made of the interaction effects of birth weight and intervention by IHDP (1990)-the "heavier" infants profited more than the "lighter" infants. We performed a regression analysis with an equation that included intervention, birth weight, and their interaction to predict child IQ. Only birth weight, $F(1, 904) = 24.5, P < .01$, and the interaction of birth weight and intervention, $F(1, 904) = 6.35, P = .012$, were statistically significant. Furthermore, the squared multiple correlation was .07-a statistically significant but small magnitude effect. This is hardly an impressive outcome, all the more so because there are other confounding factors.

Consider the issue of maternal intelligence, the variable that far and away explains most of the child IQ variance in the IHDP study. Although there is incontrovertible evidence showing a very strong familial connection between parental and child IQ, it was not until after IHDP was initiated that IQs (Peabody Picture Vocabulary Test scores) were obtained on the mothers. The original report

(IHDP, 1990) inexplicably did not include maternal IQ as a manifest variable, but did control for maternal education.

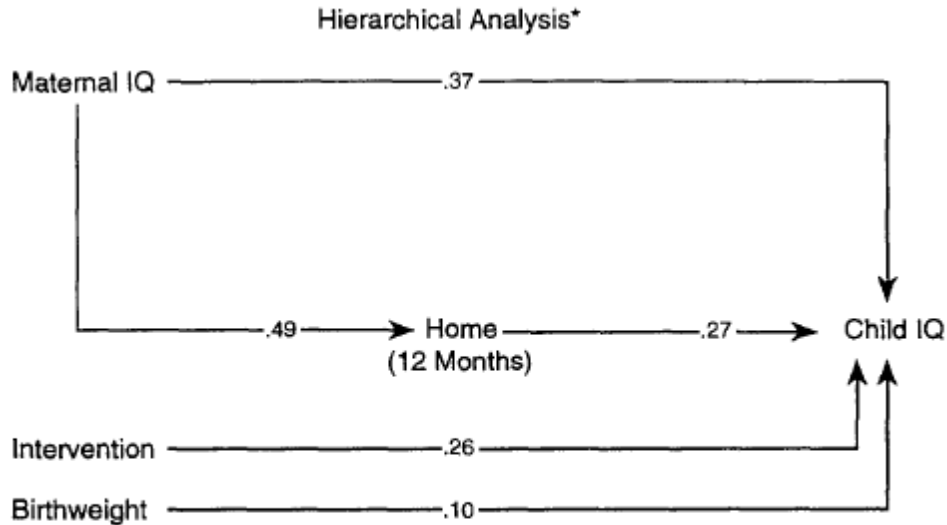
Brooks-Gunn and Benasich (1992) listed a number of valid reasons that "IHDP wanted to measure maternal IQ" including the possibility that "maternal IQ could serve as an important covariate in some statistical analyses" (p. 66). However, they went on to say that this possibility was complicated by the fact that maternal IQ was not obtained until 18 months into the program, after the onset of the intervention-"raising the possibility that maternal IQ could itself have been altered by the intervention" (p. 66). In any case, no explanation is given as to why maternal IQ was not obtained at the beginning of the study, when it could have been used as an initial status variable instead of, or along with maternal education.

Although education is often taken as a correlational proxy for intelligence, maternal education is not as powerful a predictor of child intelligence as maternal IQ. As Ramey and Ramey stated with regard to child IQ: "The single strongest predictor ... is the mother's level of tested intelligence" (1992a, p. 338).

In the case of IHDP, we found that the partial correlation for maternal education (high school or more vs less than high school) and child IQ at 36 month controlling for mother's IQ, was .09. The corresponding correlation for mother IQ, controlling for maternal education, was .43. A regression analysis including maternal IQ, intervention, birth weight, and the intervention by birth weight interaction yielded a squared multiple R of .31 ($N = 811$) which can be compared with the squared multiple R of .07 in the equation that did not include maternal IQ. In short, the comprehensive treatment package accounted for only a trivial proportion of child IQ variance. Maternal IQ overwhelmed all the other variables in predicting child IQ at 36 months.

Hierarchical Analysis

Figure 1 presents a summary of an hierarchical analysis ($N = 775$) showing partial correlation coefficients with child IQ at 36 months. The strategy for entering the variables in the equation was based on their temporal occurrence. Only statistically significant partial correlations are included in the diagram. Maternal IQ, birth weight, and intervention are treated as exogenous variables (no significant zero-order correlation was found among them). Home environment at 1 month was entered as a mediating variable.



* All coefficients are partial correlations

Figure 1. Hierarchical analysis.

Home environment did not significantly mediate, $rp = .01$, the effect of the intervention package (home visits, parents' meetings, and day-care days). But home environment contributed as much variance to child IQ (at 36 months) as the treatment. Furthermore, the home environment is a more enduring characteristic of the life experience of the child than a 2-year preschool experience. In particular, quality of caregiving has a strong transactional bearing on developmental outcome of LBW children (Sameroff & Chandler, 1975).

Another study of the mediating role of the home environment in relation to maternal and child IQ yielded essentially the same results (Bradley et al., 1993). In this case, the analysis was confined to the follow-up group and, therefore, did not incorporate the treatment aspect as we did. However, their regression analyses confirmed that maternal IQ and home environment scores, at 12 months, were a significant source of variance in child IQ scores at 36 months. Furthermore, they found that maternal IQ and home measures obtained at 36 months also contributed to child IQ. Mediated effects associated with the home measures were found at both points in time, but were stronger at 3 years.

As our hierarchical analysis reveals, maternal IQ has both a direct effect and an indirect effect on child IQ, mediated through home environment. Quite aside from the well-established genetic relationship between parental and child IQ, our analysis together with that of Bradley et al. (1993) shows that quality of caregiving is a crucial part of the linkage. This affords an opportunity for a meaningful intervention to provide stimulation and dyadic support for LBW children. That component of the IHDP intervention that consisted of home visits apparently had no effect on the home environment.

One IHDP report does show very small but significantly positive effects of maternal interactive behavior (but only for a small subset of mothers) measured

one time (in a brief, structured, and contrived laboratory assessment) on child social competence at 30 months (Spiker, Ferguson, & Brooks-Gunn, 1993). Not surprisingly, ethnicity and maternal education were significant predictors of maternal and maternal-child dyadic ratings independent of the intervention.

Intensity Effects

The claim is made that a major factor directly affecting magnitude of the intervention-determined outcome in IHDP and other child-family-oriented programs is intensity of commitment to the program (Ramey & Ramey, 1992a, 1992b). For instance, Ramey and Ramey asserted that among those families with the highest intensity of involvement, the program produced an almost "nine-fold reduction in the number of low birth weight children who were mentally retarded" (1992b, p. 133). That is a dramatic conclusion laden with important policy ramifications. As Spitz (1992) observed pursuant to his reanalysis of the Abecedarian project data: "It is no trivial matter to suggest that there may now be a method to prevent mild mental retardation, considering the long history of previous disappointments" (p. 235). But data from IHDP (1990) indicate that theirs is a very sweeping and highly misleading conclusion that must be tempered by attention to other variables such as maternal IQ and regression effects that were found when we examined data at 5 years.

Ramey et al. (1992) obtained a Family Participation Index for the intervention groups by summing number of home visits, attendance at parent group meetings, and days attended in the day-care component. The rationale presented for this summative measure was that the original intent was to evaluate efficacy of the program as a package, not as distinct components.

However, we were not similarly constrained because it is not at all clear why the intervention modalities *must* be combined and cannot be separated for purposes of analysis. Separating the intervention elements makes sense because if one modality exerts the effects, such as they are, and the others do not, how can the whole package be justified? In addition, the metric properties inherent in the Family Participation Index are highly questionable. This measure was obtained by summing the number of home visits, days the child attended the development centers, and parents' attendance at group meetings. An unwarranted assumption in creating such a summative measure is that the different components share the same scale and provide measures of a common theoretical function. The shared metric, in this case, is apparently nothing more than "frequency of contact"-a potpourri. Guralnick (1991), in reviewing the effectiveness of early intervention (including IHOP), argued that analyses of different components of comprehensive programs are necessary to identify those facets that are cost efficient. The original database breaks out the three components. We conducted a multiple regression analysis that included total numbers of center day-care days, home visits, and parent meetings, with child IQ as the dependent measure. Table 2 presents the summary of that analysis.

TABLE 2
Results of Regression Analysis Examining Relationships Between Total Day-Care Days, Meetings Attended, and Visits on Child's IQ at 36 Months

Variables in the Equation	<i>b</i>	<i>SE_b</i>	<i>B</i>	<i>r_p</i>	<i>t</i>	<i>p</i>
Meetings	1.29	.36	.23	.19	3.5	<.001
Visits	.01	.07	.01	.01	.09	.93
CDC Days	.01	.01	.04	.04	.64	.52

Adjusted $R^2 = .06$.
 $F(3,330) = 7.64, p = .0001$.

The multiple correlation was .26; the adjusted $R^2 = .06$. Again, statistically there was a small but significant effect of the entire intervention package. But the only component variable that was significant was "meetings"-neither home visitations nor the number of days attended by the child in the day program exerted an independent effect on child IQ at 36 months (see the partial correlations). It is highly improbable that parent meetings, range, 0-12, $M = 4.2, N = 334$ (those for whom child scores were available at 12 and 36 months), caused an increase in child IQ. A much more reasonable interpretation is that there was something unique about mothers who attended meetings. Children of mothers who attended the most meetings had significantly higher Bayley scores at 12 months, *prior* to the mothers having even attended any meetings, $r = .14, P < .01, N = 334$.

By far the most intensive and expensive portion of the program was the center-based day care: open for 8 hr a day, 5 days a week, 50 weeks a year, for 2 years. This is really the heart of early compensatory preschool programs, the primary motivation behind IHDP. Therefore, the day-care component bears further examination. A small zero-order correlation was obtained between number of days attended and final IQ, $r = .17, p < .01, N = 334$.

The conclusion adopted by the IHDP investigators is that those children who attended preschool more often profited in IQ terms. But a strong case can be made that some initial difference(s) distinguished children who attended day care the most frequently (and who, at 3 years, produced higher IQs) from those who attended less. Figure 2 displays child IQ at 12 and 36 months by days of attendance at the developmental centers grouped into five categories.

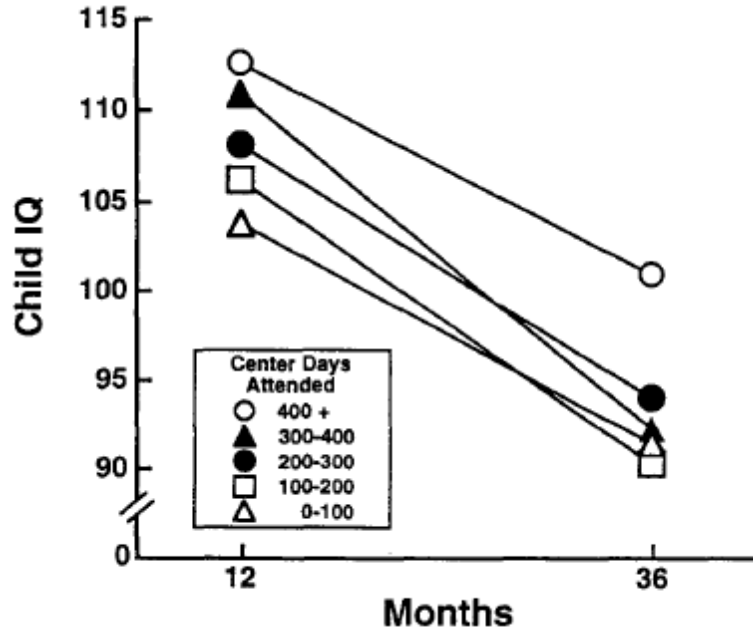


Figure 2. Child IQ by days of attendance in preschool program.

Taking the extremes of participation, 0 to 99 days ($N = 44$) versus 400+ days ($N = 69$), we found that indeed these groups differed significantly on mean IQs at 3 years, 90.3 vs 100.9, respectively, $t = 2.99$, $p < .01$. However, they also differed significantly on Bayley scores at 12 months, before the day program even began, 106.1 vs 112.5, respectively, $t = 1.91$, $P < .05$. Other categorical comparisons were also statistically significant.

There are other confounding factors that warrant consideration of the effects of the day program on IQ at 36 months. Total home visits correlated significantly with day-care days, $r = .44$, $p < .01$, $N = 334$. Could it be that home visitation prompted women with higher IQ children to send their children to the day program—children who were destined to have higher IQs at the end? On the other hand, it could be argued that day-care participation promoted home visits. However, the correlation between home visits during the first 12 months of the program (before day care started) and days attended at the developmental centers was $.27$, $p < .01$, $N = 334$, suggesting that visits affected day-care attendance, and not the other way around.

Using the extremes of participation as reflected in days attended (as already stated), those families whose children participated the least in the center day program were visited an average of 30.9 times by 12 months; those whose children attended the most were visited an average of 38.0 times by 12 months, $t = 4.2$, $p < .01$, $N = 113$. The mean IQs of the children in these groups at 36 months were 90.3 and 100.9, $t = 3.0$, $p < .01$, $N = 113$, respectively.

The only true experimental control for "participation" was the random assignment of families to intervention and follow-up groups. When data (as reported in IHDP, 1990) for the total groups at 36 months are examined, prevention of mental retardation (IQ = 70 or less, the conventional cut-off for mental retardation) is

much less dramatic than claimed by Ramey and Ramey (1992b). Of the 347 children in the intervention group tested at 36 months, 39 or 11.2% scored less than 70; for the 561 children in the follow-up group, the corresponding figures were 119 or 21.2%. Relative risk is less than 2.0. By controlling for site and initial status variables IHDP (1990) found that the adjusted odds for falling in the mental retardation range were 2.7, with a 95 % confidence interval of 1.6 to 4.8. Relatively speaking, effects of this size are still of interest, but they hardly qualify to be described as "a nine-fold reduction."

Examination of those children 1500 g or less at birth, the standard that is usually taken as an index of highly elevated risk associated with LBW, reveals that at 36 months 26.8% and 28.7% scored below 70 on the Stanford-Binet within the intervention and follow-up groups respectively (IHDP, 1990). In short, the comprehensive program had no effect on those infants at greatest risk for mental retardation.

Dividing the intervention group into participation terciles (high, medium, and low) and then comparing mean IQs, as was done by Ramey et al. (1992) without taking into account covariates, introduces all manner of confounds into the analysis. As we have shown, the components that comprise the summative Participation Index co-vary with other factors, seriously compromising the unqualified conclusion that additive "intensity" of involvement was a deciding factor in child IQ at 36 months. Initial differences explain final differences.

Conclusion

Overall, our analysis of the original IHDP data set leads to the conclusion that the reported IQ differences (and prevention of mental retardation), allegedly attributable to the intervention, rest on tenuous conceptual and methodological foundations. We have shown that the raw IQ differences (see effect sizes) between the intervention and follow-up groups as originally reported are not large and are subject to confounds. The conclusion was reached (IHDP, 1990) that "the results of this study are especially timely and relevant to the concerns at the federal and state levels for providing appropriate interventions for children at risk for developmental delay" (p. 3041). This construction of the results of the IHDP is highly questionable, not only with respect to the programmatic and policy conclusions voiced at the termination of the program, but, as importantly, whether the reported effects are enduring.

THE FIVE-YEAR FOLLOW-UP

Our independent analysis of the initial data set raises serious questions as to whether such an expensive educational intervention has any meaning even at 3 years. Now consider what happened when the children were reevaluated at 5 years, 2 years after the conclusion of the program (Brooks-Gunn, McCarton, et al., 1994).

Cognitive Outcome

The overall significant cognitive difference between intervention and follow-up groups of 9.0 Stanford-Binet points (Table 6; IHDP, 1990), at 3 years, had vanished by 5 years—mean full-scale WPPSI IQs of 92.7 and 93.2, respectively. (These means[4] were taken from the data deposited with the National Auxiliary Publication Service [NAPS], document 05156.)

For the "lighter" birth-weight infants (2000 g or less) the original effect of 7.0 points in favor of the intervention group reversed (91.1 vs 93.2). This reversal was the result of regression effects that led to a rebound by the follow-up children and to poorer performance among the intervention children. The heavier groups (2001 to 2500 g) who, because of a 13.2 IQ difference received so much attention in the original (IHDP, 1990) and in subsequent reports (e.g., Ramey & Ramey, 1992b) at 5 years tested at 96.3 and 93.3, a difference of 3.0 points, $p = .09$, in favor of the intervention group. Figure 3 presents the means for all children and for those less than 2000 g at birth at ages 3 and 5 years (from the NAPS document).

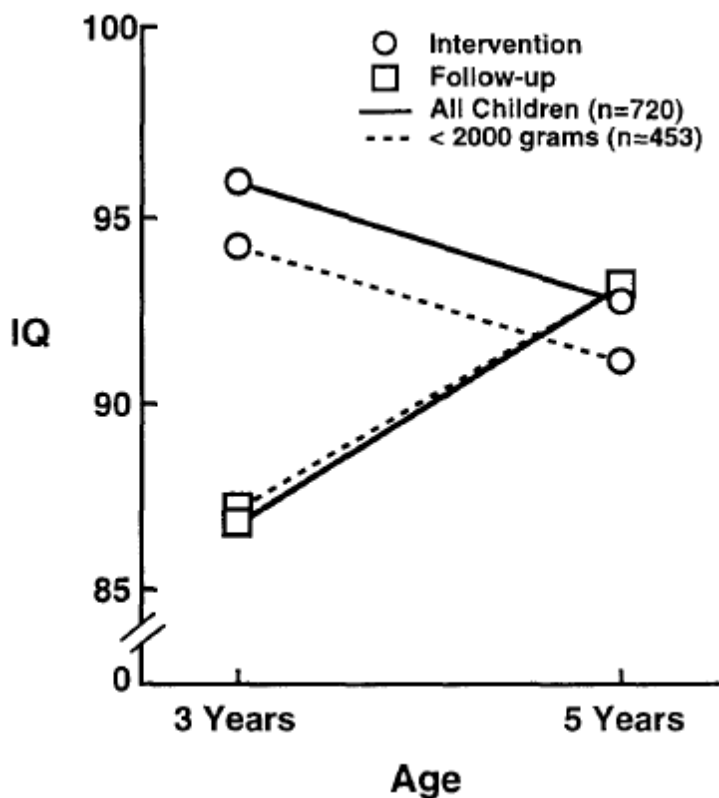


Figure 3. Child IQ at 3 and 5 years, by condition and by weight category.

Here we encounter a significant problem because not only are there major differences between the data on file with NAPS and the 1994 report by BrooksGunn, McCarton, et al., but they are of such magnitude as to reverse completely

the interpretation of the one remaining effect that Brooks-Gunn, McCarton, et al. (1994) attempted to salvage. When one compares data on deposit with the NAPS with those reported by Brooks-Gunn, McCarton, et al. (1994) there are some troubling inconsistencies.

As we have pointed out, the difference in the NAPS data for the WPPSIQs comparing the "heavier" groups is only 3.0 points, $p = .09$. Both groups showed slightly higher mean IQs than reported by Brooks-Gunn, McCarton, et al. (1994) where the difference (3.7 IQ points) reaches the conventional alpha level, $p < .05$, for rejecting the null hypothesis. In their data on file with NAPS, the difference does not reach statistical significance. Different sample sizes are reported in the two documents, although that discrepancy is not explained by Brooks-Gunn, McCarton, et al. (1994).

However, the NAPS data are superior in that they represent the same children tested at 3 and 5 years; 291 in the intervention group and 429 in the follow-up group; 73% of the total sample. Brooks-Gunn, McCarton, et al. (1994) compared different samples ($N = 312$ and 492 , respectively) and clearly included some tested at age 5 who were not tested at age 3. Those who were not tested at 3 years, but later tested at 5 years, are a biased representation of the group as a whole.

It is an interesting and rather sad commentary when the issue of effectiveness of such a large, well-publicized project boils down to quibbling over whether the one remaining effect is significant, $p = .03$, or not, $p = .09$! This discrepancy is all the more critical in that the data selected for publication not only did not square with what was reported in NAPS, but were also those data most favorable to their original hypothesis.

There is another point to be made of the comparable IQs of the intervention and follow-up children at 3 and 5 years. It was claimed that during the first 3 years the IHOP intervention "retarded" the decline in IQ observed among LBW children over the first few years of life either "temporarily or permanently" (Brooks-Gunn et al., 1993). Alas, when the children were evaluated at 5 years, just the opposite was true. The intervention children, but not the follow-up children, experienced IQ decline, as Figure 3 illustrates. [5]

Other Outcomes at Age Five. The lower incidence of behavior problems among the intervention children reported in 1990 was now absent entirely. At age 5 years the number of between-groups health conditions was no different.

COMMENT

Alleged Policy Implications

Through numerous publications and conference presentations, those associated with IHOP have emphasized important policy implications of the alleged beneficial effects of intensive and protracted intervention programs for children born at peril for adverse developmental outcomes. This is a particularly salient consideration in the case of LBW infants whom we know are at substantially elevated risk

for direct and indirect medical, educational, behavioral, and social travail. The primary rationale or premise for IHDP was that LBW infants would benefit from educational interventions that have previously been used "successfully" with children born into socioeconomically disadvantaged circumstances (BrooksGunn et al., 1993). Putting aside, for the moment, the gnawing question as to whether such a pronouncement of success of early intervention is truly justified, we focus directly on IHOP and its implications.

Our analysis has shown that the effects of intervention at 3 years of age are, in the final analysis, functionally nonexistent. In the most generous sense, IHOP was a "clean" randomized treatment study. Control for initial status variables was rationalized on the basis of conditions known to be associated with premature LBW. Testing protocols were fairly well designed and an enormous amount of data was collected.

The Weaknesses

But in another more meaningful sense the study was extremely crude and hopelessly confounded. Premature LBW «2500 g) is but a gross and imprecise risk indicator of myriad direct and indirect biological and social processes affecting the fetus and/ or mother. There are different distal and proximal etiologies of premature LBW, often appearing in conjunction-causes that have clear implications for specific intervention strategies designed either to prevent LBW or ameliorate subsequent effects on development. Furthermore, conditions that produce serious prematurity « 1500 g) are different in degree and kind from those associated with LBW in the range 2000 to 2500 g.

The premise that early educational intervention models designed for socioeconomically deprived (but NBW and full-term children) are generically appropriate for LBW infants fails to take into account etiologic specificity or related chronic outcomes-systemic and local intrauterine infections, nutrient deficiencies, gestational problems, exposure to toxins, reproductive system abnormalities, maternal health (including hypertension of pregnancy and diabetes), genetic factors, drug abuse, psychosocial stress, smoking, and others. The solution to the problem of LBW, or any health or social problem for that matter, must be measured by effectiveness with the people who are most likely to present the problem-in this case, a highly heterogeneous subpopulation of infants who are both premature and LBW.

Epidemiological and clinical studies suggest that adequacy of prenatal care, race, income, maternal age, and education, taken in conjunction with specific causative agents, structural abnormalities, respiratory and neurological deficits, specific functional limitations, and degrees of impairment are much more empirically and clinically meaningful grounds for selection of participants than gross measures of prematurity (gestational age of 37 weeks or less and birth weight of 2500 g or less). It is true that LBW infants between 1500 and 2500 g are at risk for developmental delay, possibly because of an association with adverse socioeconomic conditions. But it is the smaller babies, those under 1500 g, who present major medical problems (Alberman, 1994).

As clinical and scientific questions have become more focused and as we narrow in on causative agents, old methods of connecting crude and vague attributional factors with undifferentiated and standardized interventions, such as the IHOP, are of historical interest only, with no substance for current public health policy.

Risks conferred by birth at 35 or 36 weeks and 2000 g and above are not great, independent of other variables. Quite aside from that consideration, from an epidemiological perspective, the selection procedures employed in IHOP may have been high in sensitivity because of a net so broadly cast, but critically low in specificity. The weight gradient effect, in itself, is a critical indictment of the IHDP and its goal to reduce deleterious effects of LBW: the lighter the babies (those at greater long-term risk), the weaker the effect. At best, the results of IHDP confirm what geneticists call "reaction range" or what psychologists call the "treatment by aptitude" interaction.

The question must be raised as to whether such an expensive and ill-conceived intervention, beset with all manner of confounds, has any medical or developmental significance even at 3 years-when the measured IQ effects purportedly were greatest. There was ample reason in the literature to be suspicious about the durability of intervention effects observed at 3 years among these LBW children. For instance, a study by Wilson (1985) involving twins, one of whom was born smaller than the other, showed that LBW alone did not confer any long-term intellectual disadvantage. Although LBW was a powerful initial suppressor, as contrasted with their heavier concordant twins, LBW infants from higher SES groups recovered their IQs by the time they were in school. Wilson found what we found in the IHDP data: maternal competence is a decisive factor. This is not to say that LBW children in whom organogenesis is disrupted do not present later chronic illness quite independently of their social circumstances. The IHDP response was to exclude children who appeared to have severe illness or neurological deficits.

Predictably there was no sustained effect of the intensive program (mostly preschool education) on the sample as a whole or on the lighter «2000 g) children. Nevertheless, in the case of the "heavier" LBW babies Brooks-Gunn, McCarton, et al. (1994) claimed that the effects "are modest, but educationally relevant" (p. 1260). Even granting that an effect of 3+ IQ points was found at 5 years for the heavier children (questionable when we examine covariates at 3 years and their own data on file with NAPS), how can anyone maintain that such a trivial IQ difference has any clinical, individual meaning? IHDP was not a population-based study, so that inferences about socially valid goals based on a .25 standard deviation effect are inappropriate for LBW children weighing between 2000 and 2500 g.

Disadvantaged Children

The authors concluded that for "disadvantaged children" the intervention package is beneficial. But what is meant by disadvantage in the context of IHDP: those who fall in the weight range of 2000 to 2500 g, but not below? Or is the referent now to socially disadvantageous circumstances? The lighter babies were

not similarly "disadvantaged"? Is disadvantage the cause or the effect? The bottom line is that it is unclear that there was a meaningful intervention effect at 3 years, much less at 5 years. Claims to have prevented mental retardation were decidedly premature, turning out to be clearly erroneous. In the meantime, what happened to premature LBW and the circumstances that can avert or mitigate this and associated outcomes?

Brooks-Gunn, McCarton, et al. (1994) cautioned, as others often have before them: "The results of the IHDP call into question the notion that providing enrichment experiences in the first few years of life can protect children against biological disadvantages over extended periods of time" (p. 1261). But they went on to claim that IHDP provides evidence of "residual effects" 2 years after termination of the program. We believe the weight of evidence commands just the opposite judgment.

The results of IHDP represent the outer limit of what can be expected from early intensive intervention. Given the expense and organizational capability and professional competence required, replication of IHDP cannot be justified. Actually, there are already community programs available for these infants. At least 30% of children in the follow-up group did attend day-care programs at some time during the first 3 years (IHDP, 1990). There are so many programs that it hardly seems reasonable to add still another, particularly one that has demonstrated no specialized effect.

Other Preschool Alternatives

In fiscal years 1992 and 1993 the federal government funded 93 early childhood programs through 11 federal agencies and 20 offices (U.S. General Accounting Office, 1994). GAO identified "key" programs, programs that provide education or child care. Of these programs 22 targeted children from birth through age 5 with a budget authority of \$3.66 billion in fiscal year 1992. Economically disadvantaged children, birth through age 5, can be served by 13 programs. This is not to say that all eligible or needy children are being served. It is to say, however, that still another "key" program hardly seems necessary.

Yet, as we noted earlier, the Centers for Disease Control and Prevention are embarking on a program to provide early compensatory preschool education to at-risk children to enhance their health and development. (A detailed account can be found in Request for Proposals [RFP] 200-94-0828P, entitled "Project Begin.") The project is called Bringing Early Growth (and Development) Into Neighborhoods (BEGIN). From 1992 through fiscal year 1995 \$6.7 million was awarded for planning grants. Contracts have now been negotiated for 14 sites. BEGIN is modeled after IHDP and its predecessor, the Abecedarian project for socially disadvantaged children. A detailed and critical reanalysis of Abecedarian data by Spitz (1992, 1993) has shown claims made for the effectiveness of that project are highly questionable on methodological, analytical, and interpretative grounds.

The target population of BEGIN children are those born to women who have completed less than 12 grades of school, nothing more. (Infants will be excluded

who present serious medical complications highly associated with poor cognitive development or who have been hospitalized for 30 days or more postdelivery. Presumably these exclusion criteria would apply to many premature LBW infants.) What about social class and racial differences, differences that create groups for whom years spent in school are not qualitatively comparable? How this project is a logical extensive of IHDP (for premature, LBW children) severely challenges the imagination. What is there in the IHDP database that would suggest the early intensive preschool intervention, although demonstrably not effective for LBW infants, might be for children born to women with less than 12 grades of school (regardless of their IQ, SES, race, age, marital status, and myriad other compromising circumstances)?

CONCLUSION

Given that reanalysis of the IHDP data at 3 years produced an indeterminate outcome regarding the intensive intervention effort, that maternal competence (primarily IQ) was so overwhelming by comparison, that those babies at greatest risk < 1500 g) did not benefit at all from the intervention, that by 5 years whatever intervention effects had been initially observed dissipated completely, the IHDP should be regarded as conclusive evidence that early preschool intervention fails to alter the developmental trajectory of premature LBW babies.

Placing IHDP within the broader context of the vast research literature on preschool interventions ineluctably leads to the conclusion that effects are transitory and are completely overwhelmed by the prevailing family and community ecologies. After their analysis of the literature on this subject, Clarke and Clarke (1989) stated: "The lack of enduring benefits on cognitive skills from preschool programs has been documented so many times that it scarcely needs repeating" (pp. 291-292).

No doubt babies born too soon and too small, especially less than 35 weeks gestation and less than 1500 g, are at greatly elevated risk for adverse medical and developmental sequelae often associated with structural and pathogenic deviations. From either a clinical or a policy perspective there is much more to be gained by focus on known risk factors that are responsible, either directly or indirectly, for premature LBW. For instance, there is increasing evidence that prevention of preterm delivery caused by spontaneous preterm labor (PTL) and premature rupture of membranes (PROM), which together account for most preterm births, can be accomplished by screening and treatment for genital tract colonization with microorganisms, including pathogenic organisms in amniotic fluid and chorioamnion infection.

A recent large-scale study of over 10,000 women has shown that those diagnosed during the second trimester with bacterial vaginosis (one of the most common genital infections in pregnancy) are 40% more likely than women not infected to deliver a premature LBW baby (Hillier et al., 1995). This association holds when other known risk factors for premature LBW are controlled. Oral treatment with clindamycin has been shown to reduce by 50% preterm birth and PROM associated with bacterial vaginosis-at a cost of just a few dollars per patient (McGregor et al., 1995). Improved understanding of pathogenic pro-

cesses associated with premature delivery has led to other therapeutic interventions (Lockwood, 1994).

Primary prevention of LBW is a much more sensible and cost-efficient public health measure than treating consequences of premature delivery. In commenting on programs of early childhood interventions (specifically IHDP) Hack et al. (1994) noted that "they have little influence on structural deficits. The prevention of extreme prematurity is thus critical" (p. 758). Greater emphasis should be directed toward specific pathophysiological and biomolecular events that are linked to prematurity and LBW.

Even with much more attention and resources devoted to primary prevention of prematurity and LBW, there will inevitably be babies born small who are in need of services to counter some of the adverse consequences associated with their birth disadvantage. We have already noted that short-term interventions have produced a mix of results. But, in general, evidence of practical long-term benefits of even intensive secondary prevention measures is not reassuring-as IHDP demonstrates.

The reality is that premature LBW is a vastly complex, multifaceted, heterogeneous, and only vaguely understood biological and social constellation of causes and effects that will not yield to standardized and nonspecific interventions such as IHDP, no matter how intense and well intentioned. Proximal and distal causes are enormously diverse, as are the multiplicity of outcomes. In the aftermath of IHDP we confront a rush to judgment to provide preventive and ameliorative services without an adequate theoretical and scientific foundation. Once the developmental trajectory is set, it is extremely impervious to significant alteration by global and undifferentiated preschool interventions. Primary preventive interventions must be tailored to take into account the spectrum of biological, genetic, medical, and social factors affecting the child and the family. Demographic and psychosocial variables serve as population-based markers for identification of women who are at elevated risk for premature delivery.

NOTES

1 Funding for the IHDP was obtained from the Robert Wood Johnson Foundation, the Pew Charitable Trusts, the Bureau of Maternal and Child Health and Resources, and the National Institute of Child Health and Human Development.

2 Excluded from the original sample were 61 infants who had a severe illness or neurological defect. Of those meeting the preestablished eligibility criteria 274 (21 %) refused to consent to random assignment, leaving 1028 children. Of these, 43 withdrew before participating in the study ($N = 985$). At the conclusion of the study at 36 months 913 were available for assessment on at least one of the primary outcome measures. The primary outcome measure of greatest interest here is cognitive development. At 3 years Stanford-Binet scores were available on 908 children out of the original 985 (92% for both intervention and follow-up groups). Of course,

some attrition is to be expected, but it is uncertain what effect any loss of children at each stage had on outcome measures, particularly IQ.

3 R. T. Gross et al. Infant Health and Development Program (IHOP): Enhancing the outcomes of low birth weight, premature infants in the United States, 1985-1988 [Computer file]. Stanford, CA: Ruth T. Gross et al. (producers), 1990. Ann Arbor, MI: Inter-university Consortium for Political and Social Research (distributor), 1992.

4 These are population marginal means. Differences were derived using regression models accounting for site, gender, race/ethnicity, maternal education, maternal age, Neonatal Health Index, and birth weight.

5 Although on several occasions we requested the raw data for the 5-year follow-up we have received no response except to inform us that a "policy decision" was made not to share data until 3 years after their publication. It is our understanding that another evaluation of these children has been conducted at about age 8. Obviously these data are not available to us either.

REFERENCES

Alberman, E. (1994). Low birthweight and prematurity. In I.B. Pless (Ed.), *The Epidemiology of Childhood Disorders*. New York: Oxford University Press.

Alexander, G.R., & Korenbrot, c.c. (1995). The role of prenatal care in preventing low birth weight. *The Future of Children*, 5, 103-120.

Als, H., Lawhon, G., Duffy, F.H., McAnulty, G.B., Gibes-Grossman, R., & Blickman, J.G. (1994). Individualized developmental care for the very low birth weight preterm infant. *Journal of the American Medical Association*, 272, 853-858.

Baumeister, A.A., & Kupstas, F.D. (1990). The new morbidity: Implications for prevention and amelioration. In P.L.C. Evans & A.D.B. Clarke (Eds.), *Combatting mental handicap: A multidisciplinary approach*. Bicester, UK: AB Academic Publishers.

Baumeister, A.A., Kupstas, F.D., & Woodley-Zanthos, P. (1993). *The new morbidity: Recommendations for action and an updated guide to state planning for the prevention of mental retardation and related disabilities associated with socioeconomic conditions*. Washington, DC: U.S. Department of Health and Human Services, President's Committee on Mental Retardation. Bennett, F.C. (1987). The effectiveness of early intervention for infants at increased biological risk. In: M.I. Guralnick & F.e. Bennett (Eds.), *The Effectiveness of Early Interventions for At-Risk and Handicapped Children*. New York: Academic.

Bradley, R.H., Whiteside, L., Caldwell, B.M., Casey, P.H., Kelleher, K., Pope, S., Swanson, M., & Barrett, K. (1993). Maternal IQ, the Home Environment, and Child IQ in low birthweight, premature children. *International Journal of Behavior Development*, 16, 61-74.

Breslau, N., DelDotto, J.E., Brown, G.G., Kumar, S., Ezhuthachan, S., Hufnagle, K.G., & Peterson, E.L. (1994). A gradient relationship between low birth weight and IQ at age 6 years. *Archives of Adolescent Medicine*, 148, 377-383.

Brooks-Gunn, J., Benasich, A.A. (1992). Initial status variables. In R.T. Gross et al., *Infant Health and Development Program (IHDP): Enhancing the outcomes of low birth weight, premature infants in the United States, 1985-1988* [Computer file]. Stanford, CA: Ruth T. Gross et al., (Producers), 1990. Ann Arbor, MI: Inter-university Consortium for Political and Social Science Research (Distributor), 1992.

Brooks-Gunn, J., Gross, R.T., Kraemer, H.e., Spiker, D., & Shapiro, S. (1992). Enhancing the cognitive outcomes of low birth weight, premature infants: For whom is the intervention most effective? *Pediatrics*, 89, 1209-1215.

Brooks-Gunn, J., Klebanov, P.K., Liaw, F., & Spiker, D. (1993). Enhancing the development of low-birth-weight, premature infants: Changes in cognition and behavior over the first three years. *Child Development*, 64. 736-755.

Brooks-Gunn, J., Liaw, F., & Klebanov, P.K. (1992). Effects of early intervention on low birth weight pre term infants: What aspects of cognitive functioning are enhanced? *Journal of Pediatrics*. 120, 350-359.

Brooks-Gunn, J., McCarton, e.M., Casey, P.H., McCormick, M.C., Bauer, e.R., Bernbaum, IC., Tyson, I, Swanson, M., Bennett, F.e., Scott, D.T., Tonascia, J., & Meinert, C.L. (1994). Early intervention in low birth weight premature infants. *Journal of the American Medical Association*, 272. 1257-1262.

Brooks-Gunn, J., McCormick, M.e., Shapiro, S., Benasich, A.A., & Black, G.w. (1994). The effects of early education intervention on maternal employment, public assistance, and health insurance: The Infant Health and Development Program. *American Journal of Public Health*. 84. 924-931.

Chaikind, S., & Corman, H. (1991). The impact of low birth weight on special education costs. *Journal of Health Economics*, 10. 291-311.

Clarke, A.M., & Clarke, A.D.B. (1989). The later cognitive effects of early intervention. *Intelligence*. 13, 289-297 .

Eddy, D.M. (1994). Health system reform: Will controlling costs require rationing services? *Journal of the American Medical Association*. 272. 324-328.

Escobar, G.H., Littenberg, B., & Petitti, D.B. (1991). Outcome among surviving very low birth weight infants: A meta-analysis. *Archives of Disease in Childhood*. 66, 204-211.

Fiscella, K. (1995). Does prenatal care improve birth outcomes? A critical review. *Obstetrics & Gynecology*. 85, 468-479.

Gross, R. T. (1992). A multisite randomized intervention trial for premature, low birth weight infants: The Infant Health and Development Program. In S.L. Friedman & M. Sigman (Eds.), *The Psychological Development of Low Birthweight Children*. Norwood, NJ: Ablex.

Guralnick, M.I. (1991). The next decade of research on the effectiveness of early intervention. *Exceptional Children*. 58. 174-183.

Hack, M., Klein, N.K., & Taylor, H.G. (1995). Long-term developmental outcomes of low birth weight infants. *The Future of Children*. 5, 176-196.

Hack, M., Taylor, H.G., Klein, N., Eiben, R., Schatschneider, C., & Mercuri-Minich, N. (1994). School-age outcomes in children with birth weights under 750 g. *New England Journal of Medicine*, 331. 753-759.

Hille, E.T., Ouden, A.L.D., Bauer, L., Van Den Oudenrijn, e., Brand, R., & Verloove-Vanhorick, S.P. (1994). School performance at nine years of age in very premature and very low birthweight infants: Perinatal risk factors and predictors at five years of age. *Journal of Pediatrics*. 125(3), 426-434.

Hillier, S.L., Nugent, R.P., Eschenbach, D.A., Krohn, M.A., Gibbs, R.S., Martin, D.H., Cotch, M.F., Edelman, R., Pastorek II, I.G., Rao, A.V., McNellis, D., Regan, J.A., Carey, J.e., & Klebanoff, M.A. (1995). Association between bacterial vaginosis and preterm delivery of a low-birth-weight infant. *New England Journal of Medicine*, 333. 1737-1742.

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

Institute of Medicine, Committee to Study the Prevention of Low Birth Weight. (1985). *Preventing Low Birth Weight*. Washington, DC: National Academy Press.

Jack, B., & Culpepper, L. (1990). Preconceptual care. In I.R. Merkatz & J.E. Thompson (Eds.), *New Perspectives on Prenatal Care*. New York: Elsevier.

Jensen, A.R. (1971). Note on why genetic correlations are not squared. *Psychological Bulletin*. 75. 223-224.

Klebanov, P.K., Brooks-Gunn, 1., & McCormick, M.e. (1994). School achievement and failure in very low birthweight children. *Developmental and Behavioral Pediatrics*, 15(4), 248-256.

Klein, L., & Goldenberg, R.L. (1993). Prenatal care and its effect on preterm birth and low birth weight. In I.R. Merkatz & J.E. Thompson (Eds.), *New Perspectives on Prenatal Care*. New York: Elsevier.

Klein, N.K., Hack, M., & Breslau, N. (1989). Children who were very low birth weight: Development and academic achievement at nine years of age. *Journal of Developmental and Behavioral Pediatrics*. 10. 32-37.

Kraemer, H.C. (1988). Infant Health and Development Program Research Plan (Vol I, Section 5). In R.T. Gross et al., *Infant Health and Development Program (IHDP): Enhancing the outcomes of low birth weight, premature infants in the United States. 1985-1988 (ICPSR 9795)*. Stanford, CA: R.T. Gross et al. (Producers), 1990. Ann Arbor, MI: Inter-university Consortium for Political and Social Research (Distributor), 1992.

Kraemer, H.e., & Fendt, K. (1990). Random assignment in clinical trials: Issues in planning (Infant Health and Development Program). *Journal of Clinical Epidemiology*, 43. 1157-1167.

Lewit, E.M., Baker, L.S., Corman, H., & Shiono, P.H. (1995). The direct cost of low birth weight. *The Future of Children*, 5. 35-56.

Lockwood, e.J. (1994). Recent advances in elucidating the pathogenesis of preterm delivery, the detection of patients at risk, and preventive therapies. *Current Opinion in Obstetrics & Gynecology*, 6,7-18.

McCormick, M.C. (1989). Long-term follow-up of NICU graduates. *Journal of the American Medical Association*. 261. 1767-1772.

McCormick, M.e. (1993). Has the prevalence of handicapped infants increased with improved survival of the very low birth weight infant? *Clinics in Perinatology*. 20. 263-277.

McCormick, M.e., Brooks-Gunn, 1., Shapiro, S., Benasich, A.A., Black, G., & Gross, R.T. (1991). Health care use among young children in day care: Results in a randomized trial of early intervention. *Journal of the American Medical Association*, 265, 2212-2217.

McCormick, M.C., Brooks-Gunn, J., Workman-Daniels, K., Turner, J., & Peckham, GJ. (1992). The health and developmental status of very low-birth-weight children at school age. *Journal of the American Medical Association*. 267. 2204-2209.

McCormick, M.e., McCarton, e., Tonascia, 1., & Brooks-Gunn, J. (1993). Early educational intervention for very low birth weight infants: Results from the Infant Health and Development Program. *Journal of Pediatrics*. 123. 527-533.

McLaughlin, F.I., Altimeier, W.A., Christensen, M.1 .. Sherrod. K.B., Dietrich. M.S .• & Stern, D.T. (1992). Randomized trial of comprehensive prenatal care for low-income women. *Pediatrics*. 89. 128-132.

McGregor, J.A., French, II.. Parker. R., Draper. D., Patterson, E .. Jones. W .. Thorsgard, K., & McFee. J. (1995). Prevention of premature birth by screening and treatment for common genital tract infections: Results of a prospective controlled evaluation. *American Journal of Obstetrics and Gynecology*, 173. 157-166.

Meis, P.1., Ernest, J.M., Moore, M.L., Michielutte, R., Sharp, P .. & Buescher, P.A. (1987). Regional program for prevention of premature birth in northwestern North Carolina. *American Journal of Obstetrics and Gynecology*. 157. 550-556.

Miller, C.A., Fine, A., & Adams-Taylor, S. (1989). *Monitoring Children's Health: Key Indicators*. Washington, DC: American Public Health Association.

National Center for Health Statistics. (1993). Advance report of final natality statistics, 1990.

Monthly Vital Statistics Report. 41(9) Suppl. Hyattsville, MD: Public Health Service.

National Center for Health Statistics. (1994). Advance report of final natality statistics, 1992.

Monthly Vital Statistics Report. 43(5), Suppl. Hyattsville, MD: Public Health Service.

National Center for Health Statistics. (1996). *Health. United States*. 1995. Hyattsville, MD: Public Health Service.

Olds, D.L., Henderson, C.R., Tatelbaum, R., & Chamberlin, R. (1986). Improving the delivery of prenatal care and outcomes of pregnancy: A randomized trial of nurse home visitation. *Pediatrics*. 77. 16-8.

Ramey, C.T, Bryant. D.M., Wasik, B.H., Sparling, J.J., Fendt, K.H., & LaVange, L.M. (1992). The Infant Health and Development Program for low-birth-weight, premature infants: Program elements, family participation, and child intelligence. *Pediatrics*. 89. 454-465.

Ramey, C. T, & Ramey, S.L. (1 992a). Effective early intervention. *Mental Retardation*. 30. 337-345.

Ramey, S.L., & Ramey, C.T. (1992b). Early educational intervention with disadvantaged children—To what effect? *Applied and Preventive Psychology*, 1. 130-140.

Rauh, Y.A., Nurcombe, B., Achenbach, T, & Howell, C. (1990). The Mother-Infant Transaction Program: The content and implications of an intervention for the mothers of low birth weight infants. *Clinics in Perinatology*. 17. 31-45.

Richmond, J. (1990). Low-birth-weight infants: Can we enhance their development? *Journal of the American Medical Association*. 263. 3069-3070.

Saigal, S., Szatmari, P., Rosenbaum, P., Campbell, D., & King, S. (1991). Cognitive abilities and school performance of extremely low birth weight children and matched term control children at age eight years: A regional study. *Journal of Pediatrics*. 118. 751-760.

Sameroff, A.1., & Chandler, M.1. (1975). Reproductive risk and the continuum of caretaking casualty. In: F.D. Horowitz, M. Hetherington, S. Scarr-Salapatek. & G. Siegel (Eds.), *Review of Child Development Research* (Vol. 4). Chicago: University of Chicago Press.

Spiker, D., Ferguson, I., & Brooks-Gunn, J. (1993). Enhancing maternal interactive behavior and child social competence in low birth weight, premature infants. *Child Development*, 64, 754-768.

Spitz, H.H. (1992). Does the Carolina Abecedarian Early Intervention Project prevent sociocultural mental retardation? *Intelligence*, 16, 225-237.

Spitz, H.H. (1993). When prophecy fails: On Ramey's response to Spitz's critique of the Abecedarian project. *Intelligence*, 17, 17-23.

Tyson, I, Perlman, 1., Rosenfeld, c., & Arencibia-Mireles, O. (1992). Resource limitations: A major ethical issue in the care of indigent infants < 1500 g birth weight? *Pediatric Research*. 31,26A.

United States Department of Health and Human Services. (1989). *Caring for our future: The content of prenatal care*. (A report of the Public Health Service Expert Panel on the Content of Prenatal Care.) Washington, DC. U.S. Gov!. Printing Office.

United States General Accounting Office. (1994). *Early Childhood Programs: Multiple programs and overlapping target groups*. (GAO/HEHS 95-4FS). Washington, DC. U.S. Govt. Printing Office.

Ventura, S.I., Martin, J.A., Mathews, T.1., & Clarke, S.C. (1996). *Advance report offinal natality statistics, 1994. Monthly Vital Statistics Report*. 44(11), Suppl. Hyattsville, MD: National Center for Health Statistics.

Victorian Infant Collaborative Study Group. (1991). Eight-year outcome in infants with birthweight of 500 to 999 grams: Continuing regional study of 1979 and 1980 births. *Journal of Pediatrics*. 118. 761-767.

Wilson, R.S. (1985). Risk and resilience in early mental development. *Developmental Psychology*, 21. 795-805.

Wolke, D. (1991). Annotation: Supporting the development of low birthweight infants. *Journal of Child Psychology and Psychiatry*. 32(5). 723-741.

Zahr, L.K., Parker, S., & Cole. J. (1992). Comparing the effects of neonatal intensive care unit interventions on premature infants at different weights. *Developmental and Behavioral Pediatrics*, /3(3), 165-17 J.