The Long-Term Effects on Children and Adolescents of a Policy Providing Work Supports for Low-Income Parents

Aletha C. Huston
Anjali E. Gupta
Jessica Thornton Walker
Chantelle J. Dowsett
Sylvia R. Epps
Amy E. Imes
Vonnie C. McLoyd

Abstract

New Hope, an employment-based poverty-reduction intervention for adults evaluated in a random-assignment experimental design, had positive impacts on children's achievement and social behavior two and five years after random assignment. The question addressed in this paper was the following: Did the positive effects of New Hope on younger children diminish or even reverse when children reached the challenges of adolescence (eight years after random assignment)? Small positive impacts on school progress, school motivation, positive social behavior, child well-being, and parent control endured, but impacts on school achievement and problem behavior were no longer evident. The most likely reasons for lasting impacts were that New Hope families were slightly less likely to be poor, and children had spent more time in center-based child care and structured activities. New Hope represents a model policy that could produce modest improvements in the lives of low-income adults and children. © 2011 by the Association for Public Policy Analysis and Management.

The major purpose of this paper is to address the long-term effects on children and youth of their parents' participation in New Hope, an employment-based poverty-reduction intervention. In the 1990s, policies affecting low-income parents underwent major changes to promote work, not only by requirements that welfare recipients move toward employment, but also by increases in such work supports as the Earned Income Tax Credit (EITC), funding for child care assistance, and expanded health insurance for children. Evaluations of such policies indicate that, on average, parents' transitions from welfare into employment produced little change in overall family income and little discernible effect on children's cognitive or social development or on the immediate family environment. Programs that offered earnings supplements, health care subsidies, and expanded child care assistance, however, raised family incomes and led to positive effects on school achievement and social behavior for young children (Morris et al., 2009).

The New Hope Project was the most comprehensive of these incentive programs, providing earnings supplements and subsidies for child care and health care when participants were employed full-time, as well as community service jobs and supportive services. Its evaluation was a random assignment experiment comparing families who were offered New Hope with a control group. In follow-ups two and five years after random assignment, children in New Hope families performed better in school and evidenced more positive social behavior than did those in control families (Bos et al., 1999; Huston et al., 2003).

In several policy experiments, including New Hope, the positive effects of work-support policies on preschool and elementary school-age children did not occur for adolescents, who evidenced some increases in dropping out of school and minor deviant behavior (Gennetian et al., 2004). This pattern raises a crucial question: *Did the positive effects of New Hope and other incentive programs on younger children diminish or even reverse when children reached the challenges of adolescence?* To answer that question, we report here the results of a follow-up of the New Hope sample conducted eight years after random assignment (and five years after the program ended), when the children ranged from 9 to 19 years old.

We find no evidence of negative effects and some evidence for small lasting positive effects on school progress, school engagement and motivation, positive social behavior, and social relationships. Some of the academic and reduced problem behavior benefits observed at earlier waves were not sustained. We conclude that a policy providing basic work supports for low-income parents can modestly improve the long-term prospects of their children for staying involved in school and developing positive social skills. In separate publications, McLoyd et al. (2011) report positive impacts on adolescents' future orientation, attitudes about work, and employment behavior, especially among males, and Duncan, Huston, and Weisner (2007) report modest increases in adult employment and earnings as well as reductions in the percentage of families living in poverty.

BACKGROUND

Why New Hope Might Affect Children and Family Life

New Hope was an intervention designed to affect the work lives of adults, but there are theoretical and empirical reasons to expect that it might have indirect effects on children and family life. In the conceptual model guiding the study, the most direct effects were expected to be increases in parents' employment and earnings. Changes in economic circumstances and the resulting improved adult well-being were expected to affect home environments and positive parenting practices, and access to the New Hope subsidies for child care were expected to affect the nature and number of child care experiences, which could also lead to long-term changes in child development.

Why Age Differences in Impacts Might Be Expected

In prior evaluations, New Hope had effects on three mediators proposed in the conceptual model—parent employment, income, and formal, center-based child care. These changes in children's environments seem most likely to account for program effects on children, and each of them might affect young children's experiences differently than those of adolescents.

Employment

Among low-income families, young children of employed mothers are slightly better off than those of unemployed mothers, but virtually all of the differences appear to be accounted for by selection effects (Huston, 2002). By contrast, several longitudinal studies using a range of strong methods show negative associations between maternal employment and adolescent school performance (Gennetian, Lopoo, & London, 2008; Golden, 2005). In one study containing measures of adolescent well-being, however, mothers' entries into employment were related to improvements in adolescents' mental health, and exits from employment were associated with increased behavior problems (Chase-Lansdale et al., 2003).

Income and Poverty

Poverty during the preschool years has larger effects on later earnings and achievement than does poverty during the middle childhood and adolescent years; therefore, reducing poverty might also have more positive effects on younger children (Duncan, Ziol-Guest, & Kalil, 2010; Votruba-Drzal, 2006). Small increases in income may lead to improved resources in the home that are important to young children's cognitive development and school readiness, but less likely to change the trajectories of adolescents. Reduced poverty may also alleviate parents' stress, enabling them to be more sensitive and less harsh with their children, a benefit that would be expected to affect children's emotional well-being across the age range (Conger & Donnellan, 2007; McLoyd, 1998).

Child Care

The effects of employment may be offset by the types of settings in which children and adolescents spend time during maternal work hours. When programs help parents to use formal, center-based child care, children are provided with higher-quality care than the home-based care used by low-income families (Li Grining & Coley, 2006). High-quality center-based care is associated with school readiness and later achievement (Magnuson, Ruhm, & Waldfogel, 2007; Votruba-Drzal et al., 2010). Youth in early adolescence (about ages 11 to 13 years) show the greatest decline in achievement when parents enter employment programs (Morris, Duncan, & Clark-Kauffman, 2005); this age group may be especially vulnerable because they are too old for child care but lack the maturity and self-regulatory skills to withstand the temptations of television, deviant peers, and the like without adult supervision.

Other Potential Mediators

We also tested experimental impacts on children's household responsibilities, parents' fertility, housing mobility, health insurance, and school changes, but found no effects of New Hope (Bos et al., 1999; Huston et al., 2003); hence, these factors are unlikely to account for the program effects on children. Among mothers who had never been married at baseline, the New Hope participants (21 percent) were more likely than controls (12 percent) to be married after five years (Yoshikawa & Gassman-Pines, 2006), but there are no demonstrated benefits of marriage to a stepparent for children's development (Ginther & Pollack, 2004).

Gender Differences

The positive impacts on achievement and social behavior at both the two- and five-year assessments were greater for boys than for girls. In fact, the program increased girls' behavior problems at five years. Because control boys had considerably lower achievement and worse social behavior than control girls, the net effect of the treatment was to raise New Hope boys approximately to the levels of both program and control girls. One explanation for the gender differences arose from the ethnographic data: Parents reported investing available resources in purchases for their boys to prevent their sons from becoming involved in delinquent and criminal activities (e.g., buying an expensive pair of shoes so the boy would not steal to get them; Romich, 2009).

Why Should Effects Last Beyond the Program?

Although New Hope was not intended as a test of a time-limited program, eligibility was limited to three years because of funding constraints. Effects of early intervention

programs for children typically decline over time once the program is no longer in place, probably because many experiences and events affect intellectual and social development in ways that are unrelated to the experience of an intervention.

Program effects might be sustained or even increase if small initial differences led to sufficiently different trajectories to produce divergence over time. As economist James Heckman (2000) puts it, "success or failure [in early childhood] feeds into success or failure in school which in turn leads to success or failure in post-school learning. Early learning begets later learning and early success breeds later success just as early failure breeds later failure" (p. 5). The treatment-induced changes in environmental contexts might also continue, maintaining programmatically induced behavior. Both reduced family poverty and use of center-based care continued after the program ended (Duncan, Huston, & Weisner, 2007; Huston et al., 2005). Moreover, children and youth in New Hope families were more likely than controls to participate in such structured out-of-school activities as sports, religious groups, and clubs at all three assessments (Miller et al., 2008). Perhaps because these activities offer opportunities for supervision, mentoring, positive youth development, and contact with nondeviant peers, they promote school involvement and prosocial behavior (Mahoney, Larson, & Eccles, 2005).

Finally, as children get older, the reliability and validity of self-report measures increase, and the cross-time stability of both achievement and social behavior increases. Both of these psychometric facts reduce error of measurement, increase validity, and make it more likely that true relations of other variables to these measures will be detected.

Domains of Development

Our outcome measures include academic performance, achievement motivation, social behavior, and psychological well-being, as well as parenting and children's social relationships. Each of these domains is central to young people's successful development. Both developmental theory and economic analyses point to the joint importance of cognitive and noncognitive skills for success in the tasks of adulthood. One reason for the long-term effects of such interventions as the Perry Preschool program may be improvement in "noncognitive" skills (e.g., motivation, tenacity, perseverance, self-discipline, and social skills; Heckman & Rubinstein, 2001). In fact, early educational intervention effects on test performance typically disappear, but impacts on such indicators of minimal educational attainment as staying in grade, not being in special education, and graduating from high school more often endure (Karoly, Kilburn, & Cannon, 2005). Positive social behavior and positive relations with parents and peers are indicators of mental health in childhood and adolescence, and both are important predictors of social competence in adolescence and adulthood. By contrast, behavior problems—particularly externalizing problems, aggression, and delinquent behavior in middle childhood and adolescence—predict delinquency and aggressive disorders in adolescence and adulthood (Leschied et al., 2008; Timmermans, van Lier, & Koot, 2008).

In summary, the question addressed in this paper is the following: *Do the positive effects of New Hope on younger children continue when they reach late childhood and adolescence? That is, do these effects continue several years after the program ends?* The positive effects may dissipate or even reverse because adolescence brings increased autonomy and considerable risk for deviant behavior and school dropout that may be exacerbated by extensive parental employment. On the other hand, children whose families experienced New Hope employment-support policies in childhood may have been set on positive trajectories that are self-perpetuating; their families may have sustained some of the changes in income and child care brought about by the program; and families may have adapted to the changes

involved in increased parental employment well before children reached adolescence.

THE NEW HOPE PROJECT

We use data from the New Hope Project, an employment-based antipoverty program that provided earnings supplements and assistance with child care and health insurance expenses to low-income workers. Among the random assignment experiments testing various employment and welfare policies, Milwaukee's New Hope Project was unique because it was a community initiated, work-based antipoverty policy demonstration, not a test of welfare policies. The New Hope Project was based on two principles: (1) People who are willing to work should have the opportunity to do so, and (2) people who work full-time should not be poor. The project, conducted in two inner-city areas in Milwaukee, had four eligibility requirements: that applicants live in one of the two targeted service areas, be age 18 or older, be willing and able to work at least 30 hours per week, and have earnings at or below 150 percent of the federal poverty level.

  All participants had project representatives who provided advice, information, and job search assistance and access to a wage-paying community service job when necessary. Participants who worked full-time (30 or more hours a week) were also eligible for an earnings supplement to raise earned income above the poverty line, subsidized health insurance, and subsidized child care. Participants could use any number or combination of program benefits and services, depending on their needs. Eligibility lasted three years from the date a participant entered the program (the date of random assignment). The time limits reflected funding constraints and were not considered integral to the program's design.

  New Hope differed from Wisconsin's Aid to Families with Dependent Children (AFDC) and its system after 1996 (Wisconsin Works, or W-2) in several respects. All adults with low earnings were eligible for New Hope regardless of welfare history or family status; W-2 was available only to people with children who applied for welfare. New Hope's income threshold was 150 percent of the Federal Poverty Level (FPL); the W-2 limit was 115 percent of the FPL. Both programs offered community service jobs (CSJs) for those unable to find employment, but New Hope participants in CSJs were eligible for New Hope benefits, Earned Income Tax Credit (EITC), Social Security, and unemployment insurance credits. Wisconsin CSJs did not include any of those benefits (Kaplan & Rothe, 1999). Although participants in both programs were eligible for federal and state EITC (for non-CSJ jobs), New Hope also provided an additional earnings supplement to bring earnings above the FPL. Child care reimbursement rates were identical for the two groups, but New Hope had a higher income threshold, providing subsidies to families earning less than 200 percent of the FPL compared to 165 percent of the FPL for W-2 (Bos et al., 1999). Most W-2 participants qualified for Medicaid, but New Hope's advantage was that it provided affordable health care regardless of welfare or family status, either through HMOs used by Medicaid or through supplementing payments for employer-based insurance (Huston et al., 2003). By 1999, Wisconsin had created BadgerCare, which offered medical insurance to uninsured families with children whose incomes were less than 185 percent of the FPL (200 percent of FPL for continuing enrollees; Wolfe et al., 2006).

New Hope Evaluation Design

From August 1994 through December 1995, 1,362 adults who applied for the New Hope program and met the eligibility requirements were randomly assigned to have access to services (the program group) or to be in a control group that did not

have access to New Hope but could use any other services in the community. The experiment took place from 1994 to 1998, offering three years of eligibility for program benefits at a time when the local economy was booming (unemployment rates hovered between 3 and 5 percent). The eight-year follow-up occurred in 2003 to 2004, when the economy was slightly in decline, with local unemployment rates around 7 percent.

Of the total group, 745 adult sample members who had one or more children from the ages of 1 year and 0 months to 10 years and 11 months at the time of random assignment were selected for further study of effects on children and family life (the Child and Family Survey [CFS] sample).[1] If a family had more than one child in the target age range, two children were identified as "focal children" (n = 1,140). Sample characteristics are presented in Table 1.

Follow-up surveys conducted two, five, and eight years after random assignment included in-home interviews with parents and children as well as a mail survey to teachers. The timeline for services and data collection shown in Figure 1 lists the children's age ranges at each assessment point.

Results for the Two- and Five-Year Follow-Ups

Effects on Children's Environments

During the three years of parents' eligibility, New Hope increased the amount and stability of parental employment, increased income and reduced poverty, and increased children's time in center-based child care and out-of-school structured activities. At the five-year follow-up (two years after eligibility ended), program group parents had higher wages, and children continued to spend more time in center-based child care and structured activities (Huston et al., 2003).

Effects on Children's Development

The program had positive effects on a range of child skills and behaviors at the two-year and five-year follow-ups, when children were ages 3 to 12 and 6 to 15 years old, respectively. Program group children had higher school achievement, school motivation, positive social behavior, and fewer behavior problems than control-group children did. The positive effects occurred primarily for boys, with few positive impacts and scattered negative impacts on girls (Huston et al., 2001, 2005). By contrast, the program had slightly negative effects on the achievement of youth who were adolescents (12 or older) when it began (Bos & Michalopoulos, 2001).

Eight-Year Follow-Up Measures and Procedures

A major strength of this study is the use of multiple measures from multiple reporters to gauge child and family well-being. Administrative records provided information about quarters of employment, earnings, and welfare receipt over the eight-year period. At the three assessment points, in-person surveys with parents and children in the families' homes measured receipt of services through New Hope and other agencies; economic outcomes (e.g., hours of work, wages, and type of jobs); family functioning (e.g., well-being, parent–child relations, and discipline); children's participation in child care and out-of-school activities; and children's academic achievement, school progress, achievement motivation, positive and negative

---

[1] Those of Asian and Pacific Island descent were excluded because almost all members were Hmong refugees from Laos and Cambodia for whom the parent and child measures were deemed culturally inappropriate.

Table 1. Demographic means and *N*s imputed and unweighted.

| | Imputed | | Unimputed | | | | | | | | |
| | | | Entire Sample | | 2-year Respondents | | 5-Year Respondents | | 8-Year Respondents | |
	Mean (%)	N	Mean (%)	N	Mean (%)	N	Mean (%)	N	Mean (%)	N
Demographic controls—Family at random assignment										
Male head of household	8.48	691	8.84	656	7.84	574	8.63	556	8.24	558
Northside residence	48.44	691	48.63	656	49.13	574	49.82	556	50.18	558
Household of children and one adult	84.46	691	84.15	656	85.02	574	84.35	556	84.77	558
Household of children and two adults	15.54	691	15.85	656	14.98	574	15.65	556	15.23	558
One to two children	53.91	691	53.66	656	54.01	574	53.42	556	52.15	558
Three or more children	46.09	691	46.34	656	45.99	574	46.58	556	47.85	558
Youngest child is two years or younger	48.81	691	48.32	656	50.17	574	47.30	556	48.39	558
Demographic controls—Parent at random assignment										
Age less than 25	32.27	691	32.01	656	32.75	574	32.19	556	32.97	558
Age 25 through 34	49.59	691	49.70	656	49.65	574	48.38	556	49.28	558
Age 35 and over	18.13	691	18.29	656	17.60	574	19.42	556	7.74	558
Black	55.09	691	55.18	656	55.57	574	55.76	556	57.53	558
Hispanic	28.58	691	28.51	656	27.53	574	28.24	556	26.70	558
Non-black, non-Hispanic	16.32	691	16.31	656	16.90	574	16.01	556	15.77	558
Ever worked full-time	83.89	691	84.15	656	83.80	574	85.07	556	84.05	558
Earnings 12 months pre-New Hope $0	37.47	691	36.43	656	37.28	574	36.69	556	37.28	558
Earnings 12 months pre-New Hope $1–4,999	39.41	691	39.79	656	39.37	574	39.39	556	39.43	558
Earnings 12 months pre-New Hope $5,000 or more	23.13	691	23.78	656	23.34	574	23.92	556	23.30	558
Currently receive any government aid	82.47	691	82.01	656	82.40	574	81.12	556	82.26	558
Currently employed	38.77	691	39.78	651	40.07	569	41.37	556	40.47	556
High school diploma or GED	59.74	691	60.52	656	61.67	574	61.15	556	59.14	558
Access to a car	42.89	691	43.36	655	43.80	573	43.78	555	43.09	557
In AFDC household as a child	48.82	691	48.12	611	49.54	539	49.13	519	49.71	517
Demographic controls—Child										
Age at 5-year survey (in months)	128.50	1097	129.48	1019	129.27	907	129.74	853	129.57	869
% boys	51.82	1097	51.88	1091	51.64	916	51.68	861	51.12	941
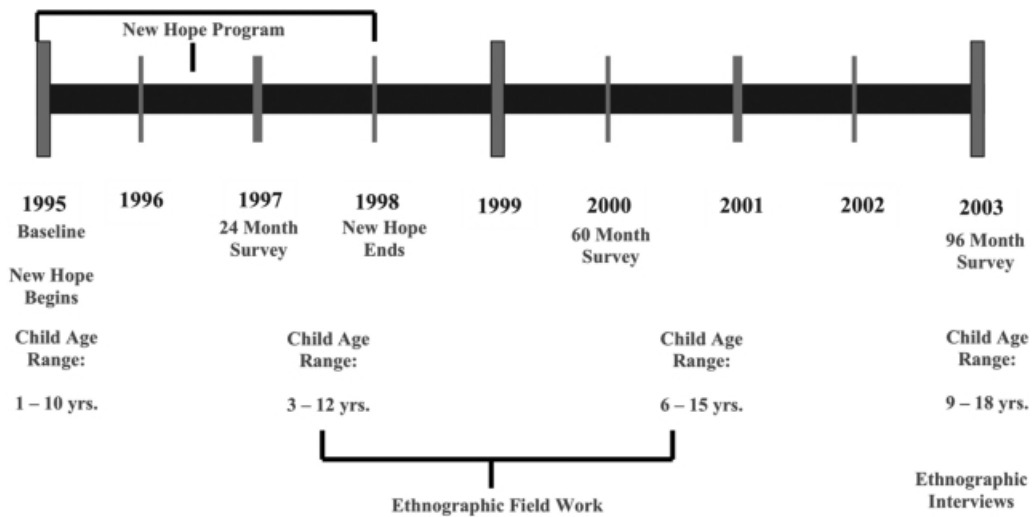
Figure 1. New Hope timeline.

social behavior, and psychological well-being. Teacher reports about children's academic performance and classroom behavior were obtained by questionnaires mailed to the children's schools. Detailed information about the measures and descriptive statistics are provided in the Appendix.[2]

Attrition

The potential sample at each assessment period consisted of all 745 families (with a total of 1,160 focal children) who were identified at baseline. At least one assessment was conducted for 691 families (93 percent of the baseline sample), all three assessments were completed with 434 families (58 percent of the baseline sample), and two of three assessments were completed with 169 families. The numbers of parents responding at each assessment were as follows: 2 years = 574 (77 percent); 5 years = 556 (75 percent); 8 years = 597 (80 percent). At eight years, parent reports were available for 938 youth (81 percent); self reports were obtained from 861 youth (74 percent). The number of responses for specific measures varied slightly.

Response rates for the mailed teacher survey were lower than those for families because teachers were contacted only after receiving parental permission, school information was sometimes incorrect or incomplete, some teachers did not return surveys, and some children were not in school. Return rates for teacher surveys for the three follow-ups were as follows: 2 years, 412 (64 percent); 5 years, 529 (48 percent); and 8 years, 544 (48 percent) of all children age 5 or older, respectively.

To evaluate possible bias resulting from attrition, we compared baseline characteristics of families surveyed and not surveyed. The eight-year responders ($n = 597$) differed significantly ($p > 0.05$) from the nonresponders ($n = 148$) on the following baseline demographic characteristics. The responders were more likely to be female, African American (and less likely to be Hispanic), to live on the north side rather than the south side of Milwaukee, to have three or more children at home, to have been receiving public assistance at baseline, and to have been in an AFDC childhood home. The two groups did not differ on experiment and control assignment, parent

age, having only one child at home, the ages of their children, amount of employment at baseline, past earnings, education, or car ownership.

Treatment of Missing Data

Missing data can lead to biased estimates of effects when the participants for whom data are missing differ systematically from those with complete data. Differential attrition reduces the ability to generalize findings to the original population. Multiple imputation is one way of protecting against such bias because it can correct for differences in the frequency of missing data for different categories of respondents. If only those cases with complete data are used, it is assumed that individuals are equally likely to have missing values on any variable. Instead, multiple imputation assumes that missing data are randomly distributed within one or more subpopulations of individuals and that plausible estimates of the missing values can be generated using observed variables in the data set.

Multiple imputation was used to replace missing data with estimates using all other observed data (including information from the two- and five-year follow-ups) to create a complete data set for the eight-year assessment. The number of data sets generated depends on the frequency of missing data (we generated ten), and the estimates of missing values within each data set differ due to a stochastic or random component that is added during the imputation process. The final analyses are performed on all data sets, generating coefficients based on a combination of the results that takes into account the variation in the estimates of the missing values (Schafer & Graham, 2002).

After eliminating the families with no data in any of the three waves ($n = 54$), we included the remaining 691 families with 1,097 children in the imputation data set. Subsequent analyses indicated virtually no differences in estimates of means or experimental impacts for cases missing one wave of data versus those missing two waves of data. For measures that should be missing—for example, measures given only to children age 12 and older—values were set to missing after the imputation. In addition, because some children with parent and child interviews were missing teacher reports, we limited the analysis sample for variables derived from the teacher survey to individuals who had received a teacher report in at least one of the three waves ($n = 863$). If teacher questionnaires were missing across all three waves prior to the imputation, values for teacher-reported child outcomes were set to missing after the imputation. The decision to delete cases after the imputation was made to reduce the amount of noise that is introduced during the imputation process when the proportion of missing data is large and identical measures at other time points are not available to improve the precision of the imputation estimates (von Hippel, 2007).

Data were imputed using the IVEware program to generate a sequential regression multivariate imputation procedure (http://www.isr.umich.edu/src/smp/ive, version dated 9/11/2006; Raghunathan et al., 2001). A normal linear regression model was used to compute missing values for all continuous variables in the imputation model. Binary variables were imputed using a logistic model, and categorical values were imputed using a polytomous or generalized logistic model. The program imputes missing values in a cyclical manner and overwrites previously drawn values to build interdependence among imputed values and exploit the correlational structure among covariates. All information across waves was used to estimate missing values, including three interaction terms: treatment*child age, treatment*gender, and treatment*prior level of earnings at baseline. The imputation model was set to use only those variables that contributed at least 1 percent of the variance to the prediction of a given missing value.

Ten data sets were created and concatenated. Because children are nested within families, parent-level variables were imputed from a family-level data file ($n = 691$),

and child-level variables were imputed from a child-level data file ($n = 1,091$). After imputation, the relevant family- and child-level variables were merged to create the final analysis data set.

As a first check on the validity of the imputation process, the means and standard errors before and after imputation were compared. These are presented in Appendix Table A.1,[3] alongside values for the non-imputed, unweighted data (i.e., raw data). In almost all cases, the imputed and original means and standard errors are quite similar. The number of nonmissing observations for each variable is shown.

Multiple imputation helps to reduce bias from attrition by using all available information to estimate values for comparable individuals, but its value depends partly on the accuracy with which the missing values can be estimated from observed values. Analyses using only complete data do not require this assumption, but they are more subject to bias if there are differences between individuals with and without missing data. Therefore, for all analyses, we present both the coefficients obtained with the raw (unweighted) data and the multiply imputed data. The point estimates from multiple imputation tended to be slightly lower than those from the raw data, but most coefficients and effect sizes were quite similar, giving us confidence that the results are robust to bias from attrition.

Analysis Model

Because New Hope was a random assignment experiment, the primary method of evaluating impacts is comparison of program and control groups. We estimated program impacts by regressing (using ordinary least squares estimation) each of our dependent measures on a dummy variable representing the program versus control group plus the following baseline covariates: high school diploma or GED, parent gender, parental age, race/ethnicity, has a child under 2 years old, has three or more children, received welfare in the prior year, AFDC in family of origin, has a car, ever been employed full-time, neighborhood, current employment status, earnings in prior 12 months, and sex and age of the focal child. Although random assignment in a large sample should ensure that the two groups do not differ on background characteristics, including these baseline covariates increases the precision of the experimental and control contrasts. The model for these regressions is shown in Equation (1).

$$Y_i = \alpha_1 + \beta_1 \text{IMPACT}_i + \beta_2 \text{COVAR}_i + e_i \tag{1}$$

Here, Y_i is the social behavior, parenting, achievement, or motivation outcome variable for child i; a_1 is a constant; $\beta_1 \text{IMPACT}_i$ is program-control difference (i.e., the experimental impact); $\beta_2 \text{COVAR}_i$ is the set of covariates listed above; and e_i is error. Two-tailed tests with an alpha of 0.10 were used. This alpha level is equivalent to a one-tailed test at $p < 0.05$, which is appropriate for the majority of program effects that were predicted, but leaves open the possibility of detecting unpredicted effects as well. The coefficients for program-control represent the difference associated with treatment.

We report effect sizes (ES) as the program-control difference divided by the standard deviation of the control group; this index indicates the effect size as a proportion of the sample standard deviation. Differences in program impacts for gender and age subgroups were tested using the HT statistic, which indicates whether the coefficients for two groups are significantly different (Greenberg, Meyer, & Wiseman, 1993). We used Stata/SE-10.1 to estimate Huber–White corrected standard errors

---

[3] All appendices are available at the end of this article as it appears in JPAM online. See the complete article at wileyonlinelibrary.com.

(White, 1982) to adjust for the fact that the observations for children within the same family are not statistically independent.

All of the analyses included the entire group of children in New Hope families and control-group families. That is, these analyses describe the intent-to-treat (ITT) or average treatment effect (ATE). The treatment was the *offer* of New Hope benefits, but, of course, not all members of the treatment group received benefits, and because the benefits were multifaceted, even those who used some of them received different components and different durations of benefits. Overall, about 87 percent of participants used at least one New Hope financial benefit. Over 80 percent received a wage supplement; over 55 percent used the health care subsidy; approximately half used the child care subsidy; and about one-third engaged in community service jobs (Huston et al., 2003).

We considered several methods of isolating the effects of receiving benefits (treatment on treated [TOT]) and for isolating the effects of different components of the treatment. This study does not satisfy the assumptions of a simple TOT correction for no-shows (Bloom, 1984; Gennetian et al., 2005). Because most people received at least some benefits, the most conservative TOT correction (no benefits vs. any) would not change impact estimates substantially. Nevertheless, to the extent that some families in the program group received minimal or no services, the true effect of the program is underestimated by these analyses. The reported coefficients could be thought of as the ITT of the actual use of services.

We also considered instrumental variables analysis (Angrist, Imbens, & Rubin, 1996) to examine the role of different treatment components (e.g., income, child care, employment). Random assignment is an ideal exogenous variable for use as an instrument, but because there is only one treatment versus control comparison in the New Hope evaluation, it cannot be used to estimate more than one mediator of treatment effects. For example, it cannot be used to estimate the separate effects of wage supplements and child care. A different instrument (e.g., a separate treatment condition) would be needed for each component (see Gennetian, Magnuson, & Morris, 2008). Variations in use of different benefits cannot be used to identify separate instruments because they are endogenous.

RESULTS

Eight years after random assignment and five years after the program ended, some of the earlier positive effects of New Hope had dissipated, but some remained or increased. Most importantly, there was no indication of reversal or negative long-term impacts on adolescents who had been in early or middle childhood when their parents entered the program. In fact, offers of New Hope's package of employment supports conferred some advantages on children as they aged into late childhood and adolescence.

The results for child measures are presented in four groups: academic achievement, achievement motivation and engagement, social behavior, and parent–child relationships. In each table, we show the impact coefficients, the control-group mean, and the effect size for the total sample. For purposes of comparison, the columns at the far right of each table indicate effect sizes from the two- and five-year assessments (from Bos et al., 1999; Huston et al., 2001, 2003, 2005). When consistent gender or age group differences in impacts exist, separate impacts for males and females or for younger (ages 9 to 12 at eight years) and older (ages 13 to 18 at eight years) youth are shown.

Academic Achievement

As shown in Table 2, the earlier program effects on most measures of achievement had dissipated after eight years. The one significant impact occurred for the aggregate

Table 2. Impacts on achievement and school progress.

Outcome	Range	Imputed Impacts				Unimputed Impacts				2-Year	5-Year
		Control Mean	Impact	SE	Effect Size[a]	Control Mean	Impact	SE	Effect Size[a]	Effect Size[a]	Effect Size[a]
Woodcock–Johnson test achievement[b]											
Broad reading score		91.86	1.45	1.02	0.10	91.74	1.68	1.15	0.11	n/a	0.12*
Male		90.98	2.37*	1.42	0.16	90.53	2.90*	1.62	0.19	n/a	0.18
Female		92.56	0.89	1.51	0.06	92.59	1.17	1.55	0.08	n/a	0.11
Applied problems		94.04	0.16	1.12	0.01	93.99	0.31	0.98	0.02	n/a	0.06
Parent ratings of achievement											
Overall achievement	1 = not well at all, 5 = very well	3.59	0.09	0.06	0.09	3.76	0.08	0.07	0.07	0.09	0.01
Literacy	1 = not well at all, 5 = very well	3.60	0.10	0.07	0.09	3.59	0.11	0.07	0.11	n/a	0.21***
Math	1 = not well at all, 5 = very well	3.58	0.08	0.07	0.07	3.58	0.10	0.08	0.09	n/a	0.05
Negative school progress	0 = no, 1 = yes	0.25	−0.03*	0.02	−0.11	0.26	−0.04*	0.02	−0.13	n/a	−0.06
Teacher ratings of achievement											
SSRS academic subscale	1 = lowest 10 percent of the class,	3.19	−0.01	0.08	−0.01	3.18	0.01	0.09	0.01	0.25**	0.06
Male	5 = highest 10 percent of the class	3.02	0.04	0.13	0.05	2.94	0.08	0.12	0.09	0.33**	0.30**
Female		3.35	−0.07	0.11	−0.07	3.42	−0.06	0.13	−0.07	0.12	−0.17
Mock report card—Reading	1 = below, 5 = excellent	2.95	−0.02	0.08	−0.02	2.95	−0.03	0.10	−0.03	n/a	0.08
Mock report card—Math	1 = below, 5 = excellent	2.78	−0.01	0.09	0.00	2.79	−0.05	0.12	−0.04	n/a	−0.02

Notes: The sample size for the imputed impactsis 1,097 (568 male; 529 female), excepting teacher-reported in pacts, which were calculated on a subset of imputed data. That subset included only data for children who had at least one completed teacher survey across the three waves (n = 8/63; average of 431.5 males and females across imputed data sets). The sample sizes for non-imputed impacts vary by reporter teacher (n = 541; 265 male, 274 female), parent (n = 926; 472 male, 454 female), and Woodcock-Johnson (n = 769; 382 male, 387 female). Actual sample size for individual measure vary as result of missing data.

[a] The effect size is the difference between program and control group outcomes as a proportion of the standard deviation of the control group. This standard deviation is always obtained from the full research sample, even if the table shows impacts for subgroups.

[b] Woodcock-Johnson scores are age-standardized with a mean of 100 and a standard deviation of 15.

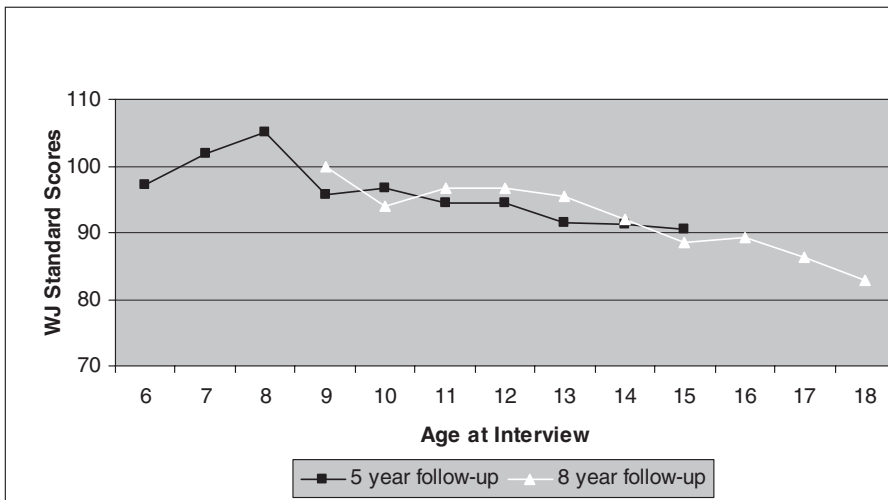Statistical significance levels are indicates *** =1 percent; ** = percent; and * = 10 percent.

Figure 2. Control group scores on Woodcock-Johnson Broad Reading scores by age.

variable labeled negative school progress ("receives poor grades," "has been in special education," and "has been retained in grade"), which represents the extent to which children fail to meet benchmarks of adequate progress in school. New Hope parents reported fewer instances of these events. Disaggregating the three components, the program effect was significant only for the single item "gets poor grades" ($\beta = -0.07$, $p < 0.05$); 30 percent of program group children and 37 percent of control-group children were described by their parents as getting poor grades.

Impacts on negative school progress appeared primarily among the children age 13 and older (under age 13, $\beta = -0.01$, n.s., ES $= -0.02$; age 13 and older, $\beta = -0.05$, $p < 0.10$, ES $= -0.16$). It is noteworthy that similar impacts were apparent at the five-year assessment for children 13 and older (Huston et al., 2003). As control-group children got older, they were increasingly likely to experience one or more of the events composing negative school progress. Although these increases also occurred for New Hope children, they were less pronounced.

None of the HT tests for gender differences in impacts on achievement were significant, but we ran separate analyses for boys and girls because of the previous findings that impacts on boys were greater than those on girls. One of these analyses replicated a pattern found at the five-year follow-up. Boys in the New Hope group scored significantly better than control boys on the Woodcock–Johnson Broad Reading scale. In Figure 2, Woodcock–Johnson reading scores for the control group in both the five- and eight-year follow-ups are shown by chronological ages of the children. At ages 6 to 8, children generally performed at or near the national average of 100; by ages 15 to 18, the average scores were less than 90, more than two-thirds of a standard deviation below the national average. Program impacts need to be considered against the backdrop of this overall tendency for scores to decline with age.

Achievement Motivation and Involvement

As shown in Table 3, compared to control-group youth, New Hope youth reported more school engagement, indicating comfort and allegiance with their schools, and higher expectations for achievement in math (unweighted impacts only). In earlier waves, New Hope boys had higher expectations about their ultimate educational

Table 3. Impacts on motivation and engagement.

Outcome	Range	Imputed Impacts Control Mean	Impact	SE	Effect Size[a]	Unimputed Impacts Control Mean	Impact	SE	Effect Size[a]	2-Year[b] Effect Size[a]	5-Year Effect Size[a]
Motivation											
School engagement	1 = none of the time, 6 = all of the time	3.71	0.15**	0.06	0.16	3.70	0.18***	0.06	0.19	n/a	0.10
Male										n/a	0.25*
Female										n/a	−0.08
English expectancy	1 = not at all well, 7 = very well	5.55	0.14*	0.08	0.12	5.54	0.16*	0.08	0.14	n/a	0.0304
Male		5.41	0.26**	0.11	0.22	5.34	0.34***	0.12	0.29	n/a	0.10
Female		5.70	0.01	0.11	0.01	5.74	−0.01	0.11	−0.01	n/a	−0.08
Math expectancy	1 = not at all well, 7 = very well	5.14	0.13	0.09	0.10	5.13	0.16*	0.09	0.12	n/a	−0.05
Male		5.20	0.13	0.13	0.10	5.24	0.12	0.13	0.09	n/a	−0.10
Female		5.08	0.13	0.13	0.10	5.02	0.18	0.13	0.14	n/a	−0.01
Certainty of educational attainment	1 = not at all sure, 5 = very sure	4.33	0.04	0.05	0.05	4.33	0.04	0.05	0.05	0.23*	0.22**
Male		4.16	0.19**	0.07	0.23	4.11	0.22***	0.08	0.27	0.46**	0.38***
Female		4.51	−0.12*	0.07	−0.15	4.56	−0.16**	0.07	−0.19	0.01	−0.06

Notes: The sample size for the imputed impactsis 1,097 (568 male; 529 female). The sample size for non-imputed impactsis 866 (436 male; 430 female). Actual sample size for individual measures vary as a result of missing data.

[a]The effect size is the difference between program and control group outcomes as a proportion of the standard deviation of the control group. This standard deviation is always obtained from the full research sample, even if the table shows impacts for subgroups.

[b]Administered to children age 9 and older.

Statistical significance levels are indicated as *** = 1 percent; **=5 percent; and* = 10 percent.

attainments and higher estimates of their competency in English than did control-group boys. That pattern continued at the eight-year follow-up. Boys in New Hope scored higher than control-group boys on expectancies for high performance in English and on their beliefs about ultimate educational attainment, replicating the earlier patterns. Program girls had lower expectations for ultimate educational attainment than did control-group girls.

Social Behavior and Parental Control

In earlier waves, there were strong positive effects on several indicators of social behavior for boys and minimal or negative effects for girls. At the eight-year follow-up, there were no effects on any teacher-reported scales, but program parents rated their children higher than did control-group parents on positive social behavior—an aggregate score composed of subscales measuring social competence, compliance, and autonomy (see Table 4). The coefficients for boys and girls were not significantly different.

Parent–Child and Peer Relationships

Parent Management

New Hope parents reported better control, fewer discipline problems, and less aggravation and parenting stress—a composite of variables that we labeled "effective child management" than did control group parents (see Table 5). In many respects, this measure can be considered another index of children's behavior problems—New Hope children behaved better and required less discipline. There were no program effects on parenting warmth (results not shown) or on children's reports of positive and negative relationships with their parents.

   Impacts on effective child management were greater for children age 13 and older than for children younger than 13, a pattern that is consistent with the five-year finding that impacts were significant for older children but not for younger children. The pattern for the control groups suggests that effective child management tends to deteriorate with age; New Hope appears to have counteracted this trend to some degree.

Peer Relations

According to children's self-reports of their peer relations, program children were less lonely and more satisfied with their friendships as compared with children in the control group (see Table 5). New Hope children also tended to have higher levels of efficacy or hope about achieving their goals (unweighted data analysis only).

CONCLUSIONS

New Hope was a comprehensive work-based poverty intervention for adults that produced improvements in children's achievement and social behavior during the program and shortly after it ended. The major question addressed in this paper was the following: Did the positive effects of New Hope on children's academic and social behavior endure when they reached adolescence? This question was especially important in light of findings from a meta-analysis of several welfare and employment experiments (including New Hope) that demonstrated *negative* effects on school performance and behavior for youth who were already adolescents when parents entered the experimental programs (Gennetian et al., 2004; Morris, Duncan, & Clark-Kauffman, 2005).

Table 4. Impacts on social behavior.

Outcome	Range	Imputed Impacts				Unimputed Impacts				2-Year	5-Year
		Control Mean	Impact	SE	Effect Size[a]	Control Mean	Impact	SE	Effect Size[a]	Effect Size[a]	Effect Size[a]
Positive behavior scale											
Parent report	1 = never, 5 = all of the time	3.77	0.08**	0.04	0.14	3.77	0.08*	0.04	0.14	0.03	0.15*
Male		3.71	0.12**	0.05	0.21	3.70	0.13**	0.06	0.23	0.22*	0.13
Female		3.83	0.02	0.05	0.04	3.85	0.01	0.05	0.02	-0.17	0.15
Teacher report	1 = never, 5 = all of the time	3.60	0.01	0.06	0.02	3.61	0.00	0.06	0.00	0.25**	0.01
Male		3.51	0.01	0.10	0.01	3.45	0.02	0.08	0.03	0.50***	0.24*
Female		3.71	-0.01	0.07	-0.01	3.76	-0.02	0.09	-0.03	0.05	-0.26**
Problem behavior scale											
Externalizing—Parent report	1 = never, 5 = all of the time	2.38	-0.04	0.05	-0.06	2.36	-0.02	0.06	-0.03	-0.02	-0.13
Male		2.45	-0.11	0.07	-0.15	2.44	-0.10	0.07	-0.14	-0.19	-0.12
Female		2.30	0.04	0.07	0.05	2.27	0.07	0.07	0.09	0.15	-0.06
Externalizing—Teacher report	1 = never, 5 = all of the time	2.06	-0.01	0.07	-0.01	2.06	-0.02	0.08	-0.02	-0.10	0.05
Male		2.14	0.00	0.12	0.01	2.18	-0.02	0.11	-0.02	-0.51***	-0.04
Female		1.98	0.00	0.10	0.00	1.95	-0.02	0.11	-0.02	0.27**	0.18
Internalizing—Parent report	1 = never, 5 = all of the time	2.37	-0.07	0.05	-0.10	2.37	-0.07	0.05	-0.11	-0.02	-0.05
Male		2.39	-0.10	0.07	-0.15	2.38	-0.10	0.07	-0.14	-0.17	-0.10
Female		2.35	-0.03	0.07	-0.04	2.35	-0.03	0.07	-0.05	0.12	-0.05
Internalizing—Teacher report	1 = never, 5 = all of the time	2.30	0.00	0.06	0.00	2.29	0.02	0.06	0.04	-0.10	0.05
Male		2.33	0.00	0.08	0.00	2.32	0.03	0.09	0.04	-0.22	-0.07
Female		2.27	0.00	0.08	0.01	2.26	0.01	0.09	0.02	0.07	0.24*
Classroom behavior scale											
Teacher report	1 = almost never, 5 = almost always	3.69	0.03	0.09	0.03	3.71	0.01	0.09	0.01	0.15	0.02
Male		3.55	-0.01	0.15	-0.01	3.44	-0.01	0.13	-0.01	0.38**	0.30**
Female		3.86	0.03	0.11	0.03	3.97	0.02	0.12	0.01	0.02	-0.27**

Notes: The sample size for the imputed impactsis 1,097 (average of 568 males and 529 female across impute data sets). Excepting teacher-reported impacts, which were calculated on a subset of imputed data. The subset included only data for children who had at least one completed teacher survey across the three waves (n = 863 average of 431 5 boys and girls across imputed data sets). The sample sizes for unimputed eight-year impacts vary by report. Teacher (n = 544; 270 boys; 274 girls), parent (n = 938; 478 male; 460 female). Actual sample size for individual measures vary as a result of missing dta.
[a]The effect size is the difference between program and control group outcomes as a proportion of the standard deviation of the control group. This standard deviation is always obtained from the full research sample, even if the table shows impacts for subgroups.
Statistical significance levels are indicated as *** = 1 percent; ** = 5 percent; and * = 10 percent.

Table 5. Impacts on social relationships.

Outcome	Range	Imputed Impacts				Unimputed Impacts				2-Year	5-Year
		Control Mean	Impact	SE	Effect Size[a]	Control Mean	Impact	SE	Effect Size[a]	Effect Size[a]	Effect Size[a]
Peer relations											
Peer relationships— Child report	1 = always true, 5 = not true at all	4.12	0.09*	0.05	0.14	4.11	0.12***	0.04	0.18	0.12	0.02
Peer conventional behaviors— Child report	1 = none of them, 5 = all of them	3.24	0.05	0.06	0.06	3.22	0.10	0.06	0.11	n/a	−0.02
Child efficacy (Hope Scale)	1 = none of the time, 6 = all of the time	3.82	0.09	0.06	0.11	3.81	0.11**	0.05	0.14	n/a	0.10
Delinquent behavior— Child report[b]	1 = never, 5 = five or more times	0.22	0.00	0.02	0.01	0.24	0.00	0.03	0.00	n/a	0.11
Effective child management[c]	1 = low, 5 = high	3.89	0.18***	0.06	0.20	3.86	0.15***	0.05	0.23	n/a	0.11
Younger[d]	1 = never, 6 = all of the time	4.03	0.07	0.08	0.08	3.96	0.04	0.07	0.07	n/a	0.17
Older[e]		3.79	0.24***	0.09	0.28	3.79	0.19***	0.07	0.30	n/a	0.32**
Problems with control		2.31	−0.22***	0.08	−0.21	2.33	−0.25***	0.08	−0.24	−0.08	−0.15*
Parenting stress	1 = not at all true, 5 = very true	1.91	−0.14**	0.06	−0.16	1.93	−0.17***	0.06	−0.21	−0.03	−0.07
Positive youth–parent relations											
Child-reported positive relations	1 = not at all true, 5 = very true	4.24	0.03	0.05	0.04	4.23	0.04	0.05	0.05	0.13	0.00
Child-reported monitoring	1 = strongly disagree, 4 = strongly agree	3.29	0.01	0.05	0.01	3.29	0.01	0.06	0.01	n/a	0.11
Negative youth–parent relations											
Child-reported negative relations	1 = not at all true, 5 = very true	2.75	−0.01	0.06	−0.01	2.75	−0.02	0.06	−0.02	0.00	0.10

Note: The sample size for the imputed impacts is 1,097. The sample sizes for non-imputed impacts vary by reporter: parent (n = 938), child (n = 862). Actual sample sizes for individual measures vary as a result of missing data.

[a] The effect size is the difference between program and control group outcomes as a proportion of the standard deviation of the control group. This standard deviation is always obtained from the full research sample, even if the table shows impacts for subgroups.

[b] Only children ages 12 and older were asked about delinquent behavior (n = 781 imputed; 611 non-imputed).

[c] Parent report.

[d] Children ages 9 to 12 at the eight-year survey (n = average of 494 children across imputed data sets; 372 non-imputed).

[e] Children ages 13 to 16 at the eight-year survey (n = average of 603 children across imputed datasets; 469 non-imputed).

Statistical significance levels are indicated as *** = 1 percent; ** = 5 percent; and * = 10 percent.

The short answer is that some positive effects endured or appeared for the first time, though others disappeared. Importantly, there is no evidence of reversal or negative impacts for this group of youth whose parents entered the program when the children were age 10 or younger. In the domain of academic performance, many of the positive effects dissipated, but there were long-term effects of the New Hope intervention on normal school progress, engagement with school, and expectancies for math performance for the total sample, and on expectancies for success in reading, expectations for educational attainment, and reading skill for boys. The effects on positive social behavior and parental control, as reported by parents, were consistent across waves, but both positive and negative impacts on teacher-reported social behavior disappeared.

The young people themselves reported better well-being on an index of loneliness and peer relationships, and they felt more optimistic and efficacious about achieving their goals. Both strong friendships and a sense of being able to reach one's goals are hallmarks of good social and mental health. The effects of strong peer attachments depend, however, on the extent to which those peers engage in prosocial or antisocial behavior. Although there were no differences in students' reports of prosocial peer behavior, we note that New Hope youth were more likely than controls to spend time in structured activities, particularly religious and volunteer activities, which are likely to provide a positive peer environment (Miller et al., 2008).

Why Some New Hope Effects Lasted

In the conceptual model we proposed that New Hope effects on children might be mediated by parent employment, reduced poverty, and children's experience in center-based child care as well as by parents' psychological well-being and parenting. One cannot disaggregate the effects of different components of the treatment while maintaining the random assignment design, but one can examine treatment effects on the proposed mediators both during and after the program to draw inferences about those that are most likely to have contributed to the impacts on children. An experimental impact on a mediator is a necessary though not sufficient condition for such an inference. Using this criterion, the most likely mediators were the increases in parents' employment, income, and use of center-based child care and structured activities observed during the three-year eligibility period. Indeed, these impacts endured to some degree well after the parents' program eligibility ended. It is particularly noteworthy that New Hope children continued to spend more time than controls did in center-based child care during the elementary years and in structured activities across all waves of data collection (Miller et al., 2008).

Effects may have endured because children's program-induced experiences occurred during "sensitive developmental periods." Income during the preschool years has stronger effects on achievement than does later family income (Duncan, Ziol-Guest, & Kalil, 2010; Votruba-Drzal, 2006), and center-based care makes its largest contribution to academic skills in the early years (National Institute of Child Health and Human Development Early Childcare Research Network & Duncan, 2003). The early advantages that accrued during New Hope's benefit period may have altered behavioral trajectories for young children as their positive behaviors elicited supportive reactions from parents, teachers, and peers—the "skills beget skills" process. We found evidence, for example, suggesting that boys' improved behavior two years after random assignment led to parents' reports of more effective discipline and management three years later (Epps & Huston, 2007).

Finally, the impacts on such noncognitive skills as school engagement, positive social skills, and peer relationships may have contributed to the small positive effects of New Hope on school progress. For children who are at high risk of school failure, school engagement and motivation, along with social and interpersonal skills, can

increase the likelihood of staying in school as well as reducing the probability of delinquent behavior (Li & Lerner, 2011). Every year in high school is associated with increased adult earnings. Specifically, for 25-year-old males, compared to dropouts with nine years of schooling, dropouts with ten years of schooling earn significantly (10 percent) more and dropouts with 11 years earn 18 percent more (Cameron & Heckman, 1993). Most of our sample was not old enough to assess high school completion or adult earnings, but boys in New Hope families were more likely to expect to attend college, had more positive attitudes about work, and were more likely to be working for pay than were control-family boys (McLoyd et al., 2011).

The experimental impacts can be best understood in the context of normative declines with age in performance, motivation, and behavior within this population. As children got older, the control group's rate of school failure increased along with declines in average reading performance, expectancies of school success, engagement in school, positive social behavior, and parent-reported effective child management. The control group represents the counterfactual (what would have happened to children in the program group without New Hope); hence, we conclude that New Hope partially counteracted these negative trends.

No Reversal in Adolescence

The generally positive or neutral effects of New Hope on youth whose parents entered New Hope prior to their adolescent years contrast with earlier findings that several welfare and employment programs, including New Hope, had negative impacts on children who were adolescents at program onset. In fact, some of the positive impacts we observed were more pronounced for the older children in our sample (13 to 18 years old) than for the younger children. Clearly, the long-term impacts for youth who experienced New Hope as children were different from the effects on those who were adolescents when the program was initiated.

Why Many Effects Dissipated

Declining experimental effects are the usual pattern for any intervention, partly because of statistical regression to the mean and partly because children (and adults) are influenced by the wide range of subsequent experiences that are unrelated to the intervention. For example, the positive effects of Head Start on preschoolers disappear when children experience poor-quality schools (Currie & Thomas, 2000). During the eligibility period, New Hope may have provided supports that helped families withstand some of the many stresses and negative life events that are common for this population. Once it ended, families may have had fewer buffers.

It is also possible that losing New Hope eligibility was stressful, leading to some negative reactions, but there is minimal support for this hypothesis. When asked about their reactions to the end of benefits, for example, 43 percent said it was not a problem, 38 percent said it was a minor problem, and 19 percent said it was a major problem (Huston et al., 2003). Given the fact that program group members had more stable employment, higher wages, and lower poverty rates than controls two years after the program ended, there is no obvious reason to conclude that the program induced long-term negative effects.

Gender Differences

The initial effects of New Hope were stronger and more positive for boys than for girls. That was still true to some degree at the eight-year follow-up, although many of the gender differences were less pronounced. We have proposed partial explanations in earlier reports (Huston et al., 2001, 2005), but the analysis of the long-term effects does little to expand them.

Methodological Caveats

Many of the effect sizes observed were small, but they are consistent with those found in other interventions. Despite the conventional wisdom, derived from Cohen (1988), that effect sizes between 0.10 and 0.30 are small, impacts of this size can have important social implications (see Ludwig & Phillips, 2007; McCartney & Rosenthal, 2000). Even small effect sizes are striking considering that the program ended five years before these data were collected. Moreover, these effect sizes represent the average treatment effect, where the treatment is the offer of benefits, not the use of them. Although the great majority (87 percent) of the sample took up some part of the offer, fewer than half used all of the benefits; therefore, it seems safe to conclude that the true magnitude of the treatment on those who received benefits was probably greater than our estimates.

Although there were significant program impacts on a number of child behaviors, we ran a large number of statistical tests, raising the issue of whether there might have been significant results on some variables by chance. We acknowledge this possibility, but we believe that our results are meaningful because they are based on a solid conceptual model, and they form consistent patterns with earlier waves. Several patterns that appeared for the older children at the five-year follow-up, for example, emerged more strongly at eight years, when more children had reached age 13 or older. We do not interpret findings on isolated variables. We tested effects on a number of variables for which impacts were never found in order to provide a fair test of the consistency of null findings as well as positive results. Moreover, if significant program effects were random, there ought to be approximately equal numbers of positive and negative impacts at the chosen alpha levels, especially given the quite reasonable hypothesis that program effects might be negative for adolescents. In fact, 23 of the 28 coefficients reported for the whole sample were in the positive direction (i.e., children in New Hope families fared better than did those in control families). With one exception for girls, no negative program effects reached a level of $p < 0.10$. Finally, even a finding of "no negative impacts" would be informative, given the earlier analyses of adolescents. It seems reasonable, therefore, to argue that the results are reliable in showing small but important lasting positive effects of New Hope.

All of the significant impacts occurred in reports by parents and youth, and almost all were in domains of positive behavior and well-being rather than problem behavior. Unlike earlier waves, there were no effects on teacher reports in any area. One reason may be that most children were in middle or high school, where teachers have much less contact with students than they do in a typical elementary school; hence, teachers may have had less basis for evaluating social behavior or even the students' overall performance across subject-matter areas.

New Hope took place in one geographic location; hence, the findings may be specific to the economic and social conditions of Wisconsin, but it is unclear whether one might expect larger or smaller effects in other locales. The state had an aggressive welfare-to-work policy along with no waiting lists for child care subsidies and medical insurance covering most children, all of which might have reduced the impact of New Hope because they were available to control families as well as New Hope families. The strong job market during much of the follow-up period could have increased the likelihood of employment for both program and control group members. Because New Hope was a volunteer program, the results are generalizable to the population of low earners who would seek the kinds of supports it offered— that is, people who think they can work full-time and are at least 18 years old.

Policy Implications

New Hope was an employment-based intervention designed to test the effects of work supports for adults. Unlike many welfare policies, one of its primary goals was

to reduce family poverty. Although its impact on family income was small, its effects on children's well-being represent an important policy outcome. Because children in low-income families are at risk of becoming the next generation of people living in poverty, any policy that can stem that tide, even modestly, deserves attention.

In many respects, New Hope was not only ahead of its time, but also has clear relevance to current social policies. Since the 1990s, federal and state governments have emphasized employment as a goal for low-income families, strengthening work supports with large increases in the EITC, increased funding for child care, and expansion of health insurance. The political discourse regarding poverty reduction, work supports, and children's development continues to revolve around the need for *affordable* medical insurance, quality child care, and earnings supplements. One example is a collection of papers on "the next generation of anti-poverty policies" (Haskins & Sawhill, 2007), in which several research-based policy proposals call for components of New Hope. Although these proposals contain several of the core elements of New Hope, they do not capture its one-stop cafeteria-style model, nor do they include subsidized community service employment or the respectful and helpful services that were integral to the New Hope philosophy (Duncan, Huston, & Weisner, 2007).

The trend toward improving work supports slowed or reversed in the early 21st century. One of the biggest gaps is child care assistance, which was guaranteed and automatic for full-time workers in New Hope. Child care funding through the Child Care and Development Fund leveled off in the mid-2000s and then began to decline, leaving long waiting lists in some states (Schulman & Blank, 2007). Given the evidence that center-based care is one means of increasing children's academic skills and school readiness, these policy changes appear especially shortsighted.

New Hope as a Prototype for Policy

Work support proposals appeal to the core American belief that work is the most legitimate route out of poverty (Gupta, Walker, & Huston, 2007), but in times of high unemployment, work supports are probably insufficient as many workers with low skills face serious difficulties in finding and maintaining full-time work. In the 2008 to 2009 recession, the rolls for cash welfare programs did not increase in conjunction with increasing unemployment (although food stamp use did increase), suggesting that the safety net for poor adults was not operating as expected (DeParle, 2009a). Scholars and commentators faced the dilemma that the U.S. has a work-based policy for the poor and little work to offer them, but acknowledged that we will not return to a pre-1996 set of policies—that is, many government supports will continue to be tied to work (De Parle, 2009b). A combination of safety net protections with job placement and community service jobs, along the lines of the New Hope model, might be especially important for maintaining well-being of children and families during such economic downturns.

Costs and Benefits of New Hope

The total cost of New Hope was approximately $6,000 per year per family in 2005 dollars (Bos et al., 2007). Although the program had a positive impact on program participants' earnings and reduced their use of welfare programs, these societal savings do not completely offset its cost. The added benefits from improved academic attainment and reduced problem behavior, especially for boys, would result in total benefits that exceed program costs, according to calculations by Bos et al. (2007). Programs that prevent school failure can also save school districts thousands of dollars per child. Other potential benefits of the program that cannot be easily valued in dollars include parents' increased self-sufficiency, improved well-being of children, and the societal value of equalizing opportunity.

ALETHA C. HUSTON is Professor of Human Development and Family Sciences, University of Texas at Austin, 1 University Station, A2702, 108 Dean Keaton, Room Seay 1.432, Austin, TX 78712-0141.

ANJALI E. GUPTA is a researcher/statistician, 5106 Ravensdale Lane, Austin, TX 78723.

JESSICA THORNTON WALKER is a doctoral candidate in Human Development and Family Sciences, University of Texas at Austin, 1 University Station, A2702, 108 Dean Keaton, Room Seay 1.432, Austin, TX 78712-0141.

CHANTELLE J. DOWSETT is a Research Associate, University of Kansas, Center for Research Methods and Data Analysis, 1425 Jayhawk Blvd., Watson Library, 470, Lawrence, KS 66045.

SYLVIA R. EPPS is a Senior Researcher, Decision Information Resources, Inc., 2600 Southwest Freeway, Suite 900, Houston, TX 77098.

AMY E. IMES is a researcher, 7309 Tanaqua Lane, Austin, TX 78739.

VONNIE C. MCLOYD is Professor in the Department of Psychology, University of Michigan, 530 Church Street, Ann Arbor, MI 48109.

REFERENCES

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. Journal of the American Statistical Association, 91, 444–455.

Asher, S. R., & Wheeler, V. A. (1985). Children's loneliness: A comparison of rejected and neglected peer status. Journal of Consulting and Clinical Psychology, 53, 500–505.

Bloom, H. S. (1984). Accounting for no-shows in experimental evaluation designs. Evaluation Review, 8, 225–246.

Bos, J. M., & Michalopoulos, C. (2001, April). Maternal employment and changes in adolescent outcomes: Evidence from two evaluations of programs that promote work. Paper presented at the biennial meeting of the Society for Research in Child Development, Minneapolis, MN.

Bos, H., Duncan, G. Gennetian, L. A., & Hill, H. D. (2007, December). New Hope: Fulfilling America's promise to "make work pay." Hamilton Project Discussion Paper No. 2007-16. Washington, DC: Brookings Institution.

Bos, J., Huston, A., Granger, R., Duncan, G., Brock, T., & McLoyd, V. (1999). New Hope for people with low incomes: Two-year results of a program to reduce poverty and reform welfare. New York: MDRC.

Cameron, S. V., & Heckman, J. J. (1993). The nonequivalence of high school equivalents. Journal of Labor Economics, 11, 1–47.

Chase-Lansdale, P. L., Moffitt, R. A., Lohman, B. J., Cherlin, A. J., Coley, R. L., Pittman, L. D., & Votruba-Drzal, E. (2003). Mothers' transitions from welfare to work and the well-being of preschoolers and adolescents. Science, 299, 1548–1552.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences, 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates.

Conger, R. D., & Donnellan, M. B. (2007). An interactionist perspective on the socioeconomic context of human development. Annual Review of Psychology, 58, 175–199.

Cook, T. D., Church, M. B., Ajanaku, S., Jr., Shadish, W. R., Jeong-Ran, K., & Cohen, R. (1996). The development of occupational aspirations and expectations among inner-city boys. Child Development, 67, 3368–3385.

Crick, N. R., & Bigbee, M. A. (1998). Relational and overt forms of peer victimization: A multi-informant approach. Journal of Consulting and Clinical Psychology, 66, 337–347.

Currie, J., & Thomas, D. (2000). School quality and the longer-term effects of Head Start. Journal of Human Resources, 35, 755–774.

DeParle, J. (2009a, February 2). Welfare aid isn't growing as economy drops off. New York Times, p. A1. Retrieved November 15, 2009, from http://www.nytimes.com/2009/02/02/us/02welfare.html.

DeParle, J. (2009b, June 1). Slumping economy tests aid system tied to jobs. New York Times, p. A10. Retrieved November 15, 2009, from http://www.nytimes.com/2009/06/01/us/politics/01poverty.html.

Duncan, G. J., Huston, A. C., & Weisner, T. S. (2007). Higher ground: New Hope for the working poor and their children. New York: Russell Sage.

Duncan, G. J., Ziol-Guest, K. M., & Kalil, A. (2010). Early-childhood poverty and adult attainment, behavior, and health. Child Development, 81, 306–325.

Eccles, J. S., & Wigfield, A. (1995). In the mind of the achiever: The structure of adolescents' academic achievement-related beliefs and self perceptions. Personality and Social Psychology Bulletin, 21, 215–225.

Epps, S. R., & Huston, A. C. (2007). Effects of a poverty intervention policy demonstration on parenting and child social competence: A test of the direction of effects. Social Science Quarterly, 88, 344–365.

Gennetian, L. A., Duncan, G., Knox, V., Vargas, W., Clark-Kauffman, E., & London, A. S. (2004). How welfare policies affect adolescents' school outcomes: A synthesis of evidence from experimental studies. Journal of Research on Adolescence, 14, 399–423.

Gennetian, L., Lopoo, L., & London, A. (2008). Maternal work hours and schooling outcomes of low-income youth in four large urban settings. Demography, 45, 31–53.

Gennetian, L. A., Magnuson, K., & Morris, P. A. (2008). From statistical associations to causation: What developmentalists can learn from instrumental variables techniques coupled with experimental data. Developmental Psychology, 44, 381–394

Gennetian, L., Morris, P., Bos, J., & Bloom, H. (2005). Coupling the nonexperimental technique of instrumental variables with experimental data to learn how programs create impacts. In H. Bloom (Ed.), Moving to the next level: Combining experimental and non-experimental methods to advance employment policy research (pp. 75–114). New York: Russell Sage Foundation.

Ginther, D. K., & Pollack, R. A. (2004). Family structure and children's educational outcomes: Blended families, stylized facts and descriptive regressions. Demography, 41, 671–696.

Golden, O. (2005). Assessing the new federalism: Eight years later. Washington, DC: Urban Institute. Retrieved from http://www.urban.org/url.cfm?ID=311198.

Greenberg, D., Meyer, R. H., & Wiseman, M. (1993). Prying the lid from the black box: Plotting evaluation strategy for welfare employment and training programs. Discussion Paper No. 999-93. Madison, WI: Institute for Research on Poverty.

Gresham, F. M., & Elliott, S. N. (1990). Social skills rating system manual. Circle Pines, MN: American Guidance Service.

Gupta, A. E., Walker, J. T., & Huston, A. C. (2007). Working families should not be poor: The New Hope program. In D. R. Crane & T. B. Heaton (Eds.), Handbook of families and poverty: Interdisciplinary perspectives (pp. 32–47). Thousand Oaks, CA: Sage.

Haskins, R., & Sawhill, I. (Eds.). (2007). The next generation of antipoverty policies. The future of Children, 17, 1–9.

Heckman, J. (2000). Policies to foster human capital. Research in Economics, 54, 3–56.

Heckman, J., & Rubinstein, Y. (2001). The importance of non-cognitive skills: Lessons from the GED testing program. American Economic Review, 91, 145–149.

Huston, A. C. (2002). Reforms and child development. The Future of Children, 12, 59–77.

Huston, A. C., Duncan, G. J., Granger, R., Bos, J., McLoyd, V. C., Mistry, R., & Ventura, A. (2001). Work-based anti-poverty programs for parents can enhance the school performance and social behavior of children. Child Development, 72, 318–336.

Huston, A. C., Duncan, G. J., McLoyd, V. C., Crosby, D. A., Ripke, M. R., Weisner, T. S., & Eldred, C. A. (2005). Impacts on children of a policy to promote employment and reduce poverty for low-income: New Hope after five years. Developmental Psychology, 41, 902–918.

Huston, A. C., Miller, C., Richburg-Hayes, L., Duncan, G. J., Eldred. C. A., & Weisner, T. S., Redcross, C. (2003). New Hope for families and children: Five-year results of a program to reduce poverty and reform welfare. New York: MDRC.

Kaplan, T., & Rothe, I. (1999). New Hope and W-2: Common challenges, different responses. Madison, WI: Institute for Research on Poverty.

Karoly, L., Kilburn, M., & Cannon, J. (2005). Early childhood interventions: Proven results, future promise. Santa Monica, CA: RAND Corporation.

Kerr, M., & Stattin, H. (2000). What parents know, how they know it, and several forms of adolescent adjustment: Further support for a reinterpretation of monitoring. Developmental Psychology, 36, 366–380.

LeBlanc, M., & Tremblay, R. (1988). A study of factors associated with the stability of hidden delinquency. International Journal of Adolescence and Youth, 1, 269–291.

Leschied, A., Chiodo, D., Nowicki, E., & Rodger, S. (2008). Childhood predictors of adult criminality: A meta-analysis drawn from the prospective longitudinal literature. Canadian Journal of Criminology and Criminal Justice, 50, 435–467.

Li, Y., & Lerner, R. M. (2011). Trajectories of school engagement during adolescence: Implications for grades, depression, delinquency and substance use. Developmental Psychology, 47, 233–247.

Li Grining, C. P., & Coley, R. L. (2006). Child care experiences in low-income communities: Developmental quality and maternal views. Early Childhood Research Quarterly, 21, 125–141.

Ludwig, J., & Phillips, D. (2007). The benefits and costs of Head Start. Social Policy Report No. 21. Ann Arbor, MI: SRCD.

Magnuson, K. A., Ruhm, C., & Waldfogel, J. (2007). Does prekindergarten improve school preparation and performance? Economics of Education Review, 26, 33–51.

Mahoney, J. L., Larson, R. W., & Eccles, J. S. (2005). Organized activities as contexts of development: Extracurricular activities, after-school and community programs. Mahwah, NJ: Lawrence Erlbaum.

McCartney, K., & Rosenthal, R. (2000). Effect size, practical importance, and social policy for children. Child Development, 71, 173–180.

McLoyd, V. C. (1998). Socioeconomic disadvantage and child development. American Psychologist, 53, 185–204.

McLoyd, V. C., Jayaratne, T. E., Ceballo, R., & Borquez, J. (1994). Unemployment and work interruption among African American single mothers: Effects on parenting and adolescent socioemotional functioning. Child Development, 65, 562–589.

McLoyd, V. C., Kaplan, R., Purtell, K. M., & Huston, A. C. (2011). Assessing the effects of a work-based antipoverty program for parents on youths' future orientation and employment experiences. Child Development, 82, 113–132.

Miller, C., Huston, A. C., Duncan, G. J., McLoyd, V. C., & Weisner, T. S. (2008). New Hope for the working poor: Effects after eight years for families and children. New York: MDRC.

Morris, P. A., Duncan, G. J., & Clark-Kauffman, E. (2005). Child well-being in an era of welfare reform: The sensitivity of transitions in development to policy change. Developmental Psychology, 41, 919–932.

Morris, P. A, Gennetian, L. A., Duncan, G. J., & Huston, A. C. (2009). How welfare policies affect child and adolescent school performance: Investigating pathways of influence with experimental data. In J. Ziliak (Ed.), Welfare reform and its long-term consequences for America's poor (pp. 255–289). New York: Cambridge University Press.

National Institute of Child Health and Human Development Early Child Care Research Network & Duncan, G. J. (2003). Modeling the impacts of child care quality on children's preschool cognitive development. Child Development, 74, 1454–1475.

Quint, J. C., Bos, J. M., & Polit, D. F. (1997). New Chance: Final report on a comprehensive program for young mothers in poverty and their children. New York: MDRC.

Raghunathan, T. E., Lepkowski, J. M., VanHoewyk, J., & Solenberger, P. (2001). A multivariate technique for multiply imputing missing values using a sequence of regression models. Survey Methodology, 27, 85–95.

Romich, J. (2009). Trying to keep children out of trouble: Child characteristics, neighborhood quality, and within-household resource allocation. Children and Youth Services Review, 31, 338–345.

Schafer J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. Psychological Methods, 7, 147–177.

Schulman, K., & Blank, H. (2007, September). State child care assistance policies 2007: Some steps forward, more progress needed. Issue Brief. Washington, DC: National Women's Law Center.

Snyder, C. R., Sympson, S. C., Ybasco, F. C., Borders, T. F., Babyak, M. A., & Higgins, R. L. (1996). Development and validation of the State Hope Scale. Journal of Personality and Social Psychology, 70, 321–335.

Statistics Canada. (1995). Self-sufficiency project: Self complete questionnaire, parents. Montreal: Statistics Canada.

Steinberg, L., Lamborn, S. D., Dornbusch, S. M., & Darling, N. (1992). Impact of parenting practices on adolescent achievement: Authoritative parenting, school involvement, and encouragement to succeed. Child Development, 65, 1266–1281.

Timmermans, M., van Lier, P. A. C., & Koot, H. M. (2008). Which forms of child/adolescent externalizing behaviors account for late adolescent risky sexual behavior and substance use? Journal of Child Psychology and Psychiatry, 49, 386–394.

von Hippel, P. T. (2007). Regression with missing Ys: An improved strategy for analyzing multiply-imputed data. Sociological Methodology, 37, 83–117.

Votruba-Drzal, E. (2006). Economic disparities in middle childhood development: Does income matter? Developmental Psychology, 42, 1154–1167.

Votruba-Drzal, E., Coley, R. L., Maldonado- Carreño, C., Li-Grining, C. P., & Chase-Lansdale, P. L. (2010). Child care and the development of behavior problems among economically disadvantaged children in middle childhood. Child Development, 81, 1460–1474.

White, H. (1982). Maximum likelihood estimation of misspecified models. Econometrica, 50, 1–25.

Wolfe, B., Haveman, R., Kaplan, T., & Cho, Y. (2006). SCHIP expansion and parental coverage: An evaluation of Wisconsin's BadgerCare. Journal of Health Economics, 25, 1170–1192.

Woodcock, R. W., & Johnson, M. B., (1990). Woodcock-Johnson psycho-educational battery—revised. Allen, TX: DLM Teaching Resources.

Wright, J. C., & Huston, A. C. (1995). Effects of educational TV viewing of lower income preschoolers on academic skills, school readiness, and school adjustment one to three years later. Lawrence, KS: Center for Research on the Influences of Television on Children.

Yoshikawa, H., & Gassman-Pines, A. (2006). Five-year effects of an anti-poverty program on marriage among never-married mothers. Journal of Policy Analysis and Management, 25, 11–30.

APPENDIX

Descriptions of Measures at Eight-Year Follow-Up

The measures used to gauge child and family well-being are described below. Descriptive statistics are presented in Table A.1.

ACADEMIC ACHIEVEMENT AND ACHIEVEMENT MOTIVATION MEASURES

Having multiple sources of data from child, teacher, and parent reports brings strength to the evaluation. Findings have higher validity when based on several sources. A source outside of the family, such as a teacher, or an objective measure, such as a standardized test, makes findings more robust. Therefore, inclusion of multiple information sources—standardized achievement test scores, parents, teachers, and children's reports—provides a clearer picture of children's academic achievement than any one of them would alone.

Standardized Achievement Test Scores

To assess reading and mathematical competencies, children completed three individually administered scales from the Woodcock–Johnson Achievement Battery (Woodcock & Johnson, 1990). Two of these (Letter-Word Identification and Passage Comprehension) measure reading skills; the average of these two is the Broad Reading score. The third, the Applied Problems scale, measures mathematics skills. The Woodcock–Johnson was selected because its normative sample is large and representative and because it includes children from diverse ethnic groups and diverse types of schooling. The standard score for each scale is obtained by comparing the child's score with norms for his or her chronological age group. The mean standard score for the U.S. population is 100; the standard deviation is 15.

Parent Reports of Achievement

Parents' ratings of *overall achievement* were obtained from a question asking about their children's general level of achievement, based on their knowledge of their children's school progress over the past year, using a 5-point scale (1 = "not at all well," 5 = "very well"). Using the same 5-point scales, parents also evaluated their children's performance in reading and written work over the past year, which were averaged to form a *literacy scale*, and in *mathematics*.

   Parents also rated children's school progress, responding "yes" or "no" to a set of three questions about positive school progress over the past year (whether the child had been in a gifted program, received school awards for academic achievements, or awards for other types of achievement) and a set of three questions about *negative school progress* over the past year (whether the child had been in special education, repeated a grade, or received poor grades).

Teacher Reports of Achievement

The teacher survey included the *academic subscale* of the Social Skills Rating System (SSRS) (Gresham & Elliot, 1990). On this ten-item measure, teachers rated children's performance in comparison to others in the same classroom on reading skill, math skill, intellectual functioning, motivation, oral communication, classroom behavior, and parental encouragement, using 5-point scales (1 = "lowest 10 percent of the class," 5 = "highest 10 percent of the class").

Table A.1. Means, SD, and *N* of all constructs.

		Imputed[a]			Unimputed[a]		
		Mean	SD	N^b	Mean	SD	N
Woodcock–Johnson test of achievement							
Broad reading score		92.58	14.52		92.57	14.50	770
Applied problems		94.12	13.13		94.14	13.10	766
Parent ratings of achievement							
Overall achievement	1 = not well at all, 5 = very well	3.63	0.98		3.80	1.10	909
Literacy	1 = not well at all, 5 = very well	3.65	1.03		3.64	1.03	909
Math	1 = not well at all, 5 = very well	3.62	1.09		3.62	1.09	909
Negative school progress	0 = no, 1 = yes	0.24	0.30		0.24	0.30	930
Teacher ratings of achievement—Run for smaller sample							
SSRS academic subscale	1 = lowest 10 percent, 5 = highest 10 percent (of class)	3.18	0.94	863	3.18	0.97	540
Mock report card—Reading	1 = below, 5 = excellent	2.94	1.04	863	2.93	1.06	518
Mock report card—Math	1 = below, 5 = excellent	2.78	1.13	863	2.77	1.17	413
Motivation							
School engagement	1 = none of the time, 6 = all of the time	3.79	0.93		3.79	0.93	852
English expectancy	1 = not well at all, 7 = very well	5.62	1.16		5.62	1.16	783
Math expectancy	1 = not well at all, 7 = very well	5.20	1.31		5.20	1.32	784
Certainty of educational attainment	1 = not at all sure, 5 = very sure	4.35	0.78		4.35	0.80	867
Positive behavior scale							
Parent report	1 = never; 5 = all of the time	3.80	0.56		3.81	0.56	910
Teacher report[c]	1 = never; 5 = all of the time	3.61	0.68	863	3.61	0.68	545
Problem behavior scale							
Externalizing—Parent	1 = never; 5 = all of the time	2.36	0.75		2.35	0.75	910
Externalizing—Teacher[c]	1 = never; 5 = all of the time	2.06	0.88	863	2.06	0.87	546
Internalizing—Parent	1 = never; 5 = all of the time	2.33	0.67		2.34	0.67	913
Internalizing—Teacher[c]	1 = never; 5 = all of the time	2.30	0.68	863	2.30	0.68	538
Classroom behavior scale—Teacher	1 = almost never, 5 = almost always	3.71	1.01	863	3.71	1.01	543

(Continued)

Table A.1. (*Continued*)

		Imputed[a]			Unimputed[a]		
		Mean	SD	N[b]	Mean	SD	N
Effective child management	1 = low, 5 = high	3.98	0.84		3.93	0.65	825
Problems with control	1 = never, 6 = all of the time	2.20	1.02		2.20	1.02	906
Parenting stress	1 = not at all true, 5 = very true	1.84	0.83		1.84	0.83	828
Positive youth–parent relations							
Child-reported positive relations	1 = not at all true, 5 = very true	4.26	0.72		4.25	0.72	861
Child-reported monitoring	1 = strongly disagree, 4 = strongly agree	3.29	0.77		3.29	0.77	781
Negative youth–parent relations							
Child-reported negative relations	1 = not at all true, 5 = very true	2.74	0.81		2.74	0.81	860
Social relationships							
Peer relationships—Child	1 = always true, 5 = not true at all	4.16	0.65		4.17	0.64	862
Peer conventional behaviors—Child	1 = none of them, 5 = all of them	3.27	0.87		3.27	0.87	784
Child efficacy (Hope Scale)	1 = none of the time, 6 = all of the time	3.87	0.78		3.86	0.77	858
Delinquent behavior—Child[d]	1 = never, 5 = five or more times	0.24	0.30	781	0.24	0.30	611

[a] The minimum and maximums for the original (non-imputed) data show the range of individual scores. The means and standard errors shown for the imputed variables are the averages across the ten data sets, the minimums and maximums are the range of means (not individual scores) for the ten data sets.

[b] Unless otherwise noted, the N is 1,097.

[c] Teacher-reported impacts (imputed only) were calculated on a subset of data for children who had at least one completed teacher survey across the three waves.

A *mock report card* completed by teachers indicated children's current school performance in reading and math, using 5-point scales from "below" (well below grade level) to "excellent" (well above grade level). This measure was adapted from one used in the NICHD Study of Early Child Care and Youth Development.[4]

Teachers also completed the *Classroom Behavior Scale*, which contains items concerning children's study skills, conformity to classroom rules and routines, ability to work and complete tasks independently, and ability to make transitions without becoming distracted (Wright & Huston, 1995). Teachers rated children using scales from "almost never" to "almost always."

Children's Reports of Achievement Motivation

Children indicated their comfort and allegiance with their school (i.e., *school engagement*), using a 5-point scale (1 = "not true," 5 = "always true").[5] In this 5-item measure, children endorsed statements such as "you feel like you are a part of your school" and "the teachers at your school treat students fairly."

Children reported on their expectations for their performance in English and math (*English expectancy, math expectancy*), using 7-point scales (1 = "not at all well," 7 = "very well"; Eccles & Wigfield, 1995; adapted from the Self- and Task-Perception Questionnaire). In each of these subject areas, children answered questions regarding their self-concept of ability, expectations for success, extrinsic and intrinsic utility value, and attainment value.

Children were asked to indicate their educational expectations—how sure they were that they would finish high school, go to college, and finish college—using 5-point scales (1 = "not at all sure," 5 = "very sure"; Cook et al., 1996). These were averaged to form the measure "certainty of educational attainment."

SOCIAL BEHAVIOR MEASURES

Positive Social Behavior

Most studies of children from low-income families emphasize the negative aspects of social behavior. This study gives equal emphasis to positive and problem behavior. Both parents and teachers completed the Positive Behavior Scale (Quint, Bos, & Polit, 1997). Its 25 items are divided into three subscales: compliance and self-control (e.g., "Thinks before he/she acts," "Usually does what I tell him/her"); social competence and sensitivity ("Gets along well with other children," "Shows concern for other people's feelings"); and autonomy ("Tries to do things for him/herself," "Is self-reliant"). Both parents and teachers completed these scales.

Problem Behavior

Both parents and teachers rated children on *externalizing* and *internalizing* problems, using the Problem Behavior Scale of the Social Skills Rating System (Gresham & Elliott, 1990). Externalizing problems include aggression and lack of behavior control (e.g., "Is aggressive toward people or objects," "Has temper tantrums"). Internalizing problems include social withdrawal and excessive fearfulness ("Appears lonely," "Acts sad or depressed").

SOCIAL RELATIONSHIPS

Measures of parenting including parent reports and child reports were among the measures of social relationships. Most parenting measures were grouped into four

---

[4] This measure can be found at http://secc.rti.org.
[5] Items from ADD Health (http://www.cpc.unc.edu/projects/addhealth/).

composite scores: *effective child management, positive youth–parent relations, negative youth–parent relations*, and *warm and structured parenting*. These groupings were formed partly on the basis of a factor analysis of all the parenting measures at the five-year follow-up, which indicated that the sets of measures grouped on four factors.[6] These composites are likely to be more reliable than the individual scales composing them because they contain more items.

Effective Child Management

The composite variable "effective child management" represented high control (that is, few problems), infrequent discipline or punishment, low parenting stress, and high confidence in the ability to prevent harm. *Problems with control* was assessed using a five-item scale from the Canadian evaluation of the Self-Sufficiency Project (SSP) describing the frequency with which the child ignored or failed to obey the parent (Statistics Canada, 1995). Using a 6-point scale, ranging from "never" to "all of the time," parents were asked to indicate the frequency of five events (e.g., how often the child ignores the parent's punishment). The scale had a reliability coefficient of 0.80 (Statistics Canada, 1995). *Frequency of discipline* involved six items assessing the frequency, in the prior week, with which parents had punished the child by grounding, taking away privileges, and spanking (Statistics Canada, 1995). *Parenting stress* included five questions concerning the degree of difficulty that parents experienced interacting with and caring for their children (Quint, Bos, & Polit, 1997). *Confidence in preventing harm* was assessed with a single item from the parent interview: "How confident are you that you will be able to prevent your child from getting into trouble?"

Positive Youth–Parent Relations

The composite "positive youth–parent relations" was based on three child report measures: high positive parent–youth relations, high parental acceptance and involvement, and high monitoring. Children's perceptions of *positive relations* were assessed by the Child Evaluation of Relationship with Mother/Caregiver (McLoyd et al., 1994). The *acceptance/involvement* subscale of the Authoritative Parenting Measure assessed the youths' perceptions that parents were supportive and involved in their lives (Steinberg et al., 1992). Children's reports of *parental monitoring* were measured by asking children about the extent to which their parents knew about their activities and their friends. The fact that this scale correlates with other indicators of positive parent–child relations (from the child's point of view) is consistent with recent evidence that "monitoring" is an index of children's willingness to communicate with parents as well as parental efforts at supervision (Kerr & Stattin, 2000).

Negative Youth–Parent Relations

The composite variable "negative youth–parent relations" was created from two child-report variables: the "negative relations scale" from the Child Evaluation of Relationship with Mother/Caregiver (McLoyd et al., 1994) and low scores on "psychological autonomy granting" from the Authoritative Parenting Measure (Steinberg et al., 1992).

Peer Relations

The Loneliness and Social Dissatisfaction Questionnaire is a 16-item scale measuring the child's satisfaction with peer relations and friendships (Asher & Wheeler,

---

[6] A factor analysis of the measures produced four factors, which accounted for 76 percent of the variation. Measures with factor loadings higher than 0.45 were summed to form four composite scores. Details available upon request.

1985). The items are statements (e.g., "It's hard for me to make new friends"). The child answers on a 5-point scale, ranging from 1 ("always true") to 5 ("not true at all"). For this study, high scores indicate satisfaction with friendships.

Hostile Intent Attribution

This measure is designed to measure children's aggressive tendencies. It consists of four vignettes presenting situations in which another person does something that could be perceived as hostile (e.g., "A kid spilled milk down your back while you were sitting in the school cafeteria"). Respondents are asked why the person did this. The choice of answers includes benign intent (such as "The kid slipped on something") or hostile intent ("The kid wanted to make fun of you"). Two of the stories involve physical hostile intent, and two involve social hostility (e.g., not inviting someone to a party). Other research indicates that aggressive children are more likely to attribute hostile intent to others; nonaggressive children are more likely to attribute benign intent. Girls are more apt to engage in social aggression, and boys are more apt to engage in physical aggression (Crick & Bigbee, 1998).

Peer Conventional Behavior

A measure of Peer Group Conventional Behaviors, taken from the Pittsburgh Youth Study, was used to assess how many of the respondent's close friends got good grades and participated in sports, school activities, and religious activities.

Efficacy (Hope)

Children's sense of efficacy was measured using six items from the Children's Hope Scale (Snyder et al., 1996). Each subscale includes three items using a 6-point scale ranging from "none of the time" to "all of the time." Sample items include: "I think I'm doing pretty well" and "Even when others want to quit, I know I can find ways to solve the problem."

Trouble Index

Parents responded "yes = 1" or "no = 0" to items asking if children had been suspended from school, involved with police or illegal behavior, or involved with drugs or alcohol. The trouble index was the average number of yeses.

Delinquent Behavior

Administered to children age 12 and older, this scale contained 15 items adapted from LeBlanc and Tremblay's (1988) 27-item measure assessing adolescents' self-reported deviant behavior. Youth responded on a 5-point scale ranging from "never" to "5 or more times" to answer such questions as "During the past 12 months did you take part in a gang fight?" There are four subscales: fighting, stealing, vandalism, and drugs. The total scale was validated on more than 6,000 teenagers in Quebec (alpha = 0.66; LeBlanc & Tremblay, 1988).