

Helping Others Pursue Excellence in Public Schools: Assessing the Impact of HOPE CDC's Mentoring Program

Stephen M. Haas, Ph.D., Director

Erica Turley, B.S., Research Analyst

While mentoring programs are generally well received interventions, research indicates mixed results in terms of their impact. Nearly everyone is familiar with the Big Brothers Big Sisters of America organization and likely has positive attitudes toward their programs which started forming as early as 1904. Success of mentoring programs is, however, often contingent upon the program design and implementation. Those programs that have had proven impacts on relationships, attitudes, school attendance and performance, and anti-social behaviors are well designed and follow strict implementation models. Little is also known about whether these impacts, when found, can last over time. Follow-ups, when available, are generally short term and not all that favorable.

This report illustrates the results of an impact study of the HOPE (Helping Others Pursue Excellence) Community Development Corporation's mentoring program. The program is a faith-based initiative designed to improve academic performance and behavior of at-risk youth by providing mentors in typically under performing schools. The mentors or Youth Development Specialists, seek to develop positive relationships with the youth by engaging in various activities mainly at the school. Academic tutoring and lessons related to moral development are also provided to encourage the youth to become better students and citizens and ultimately prevent delinquency.

In a prior study (Haas and Turley, 2008), characteristics of the design and implementation of the HOPE CDC's mentoring program were measured against those of programs

that have proven successful in the past. The goal of that study was to determine if the program contained elements that would suggest the potential for the positive impacts they envisioned. While the HOPE CDC's program was generally well received by those involved, the study found problems with its design and implementation strategy that could hinder its ability to produce positive impacts.

Like other newly developed prevention and intervention programs, HOPE CDC experienced common implementation issues at the beginning of the school year. In addition, the study found that the model chosen as a basis for the HOPE CDC mentoring program was not evidence-based and that they departed from this curriculum in several important ways. For example, the mentors were found to be managing large caseloads rather than developing close personal relationships. Both program and school staff indicated that there was a heavy focus on academic performance and tutoring rather than mentoring. Other weaknesses of the program included little or no evidence of formal performance monitoring and an inadequate use of community resources, including family members.

Despite these programmatic issues, school staff indicated a high level of support for the program and its expansion. For the most part, the school staff reported positive relationships with the mentors and wished there were more of them. Program staff also had a genuine interest in helping as many students as possible. Finally, the students seemed to be encouraged by the program's use of incentives for good behavior and performance.

The methodology employed by this impact evaluation is supported by prior research studies involving mentoring programs. Background information and the results of some of these studies will be presented first. A brief description of the HOPE CDC mentoring program will then be provided.

Impact Evaluation Methodologies

While school-based mentoring programs continue to grow at a fast pace, the published evidence regarding their effectiveness still varies widely (Karcher, 2008; Portwood, Ayers, Kinnison, Waris, and Wise, 2005; Karcher, Kuperminc, Portwood, and Sipe, 2006). Evidence of some positive effects is more readily available for community-based mentoring programs (DuBois, Holloway, Valentine, and Cooper, 2002; Grossman and Rhodes, 2002; Tierney, Grossman, and Resch, 1995). In an experimental impact study of local Big Brothers Big Sisters affiliates, Tierney, Grossman, and Resch (1995) provided what has become perhaps the most widely cited evidence that community-based mentoring programs can have positive effects. However, experts in the field continue to call for rigorous impact evaluations specific to the newer school-based models. As resources become scarce, it is even more important to understand what types of programs work best and for which types of students so funds can be allocated appropriately (Karcher, 2008; Rhodes and DuBois, 2006; Wheeler, Keller, and DuBois, 2010; MENTOR, 2003).

The current study, therefore, followed these recommendations and utilized the most rigorous of research methods, the experimental design, with both pretesting and posttesting of randomly assigned treatment and control groups. Too few prior research studies in the area of school-based mentoring have attempted this goal. Many evaluation studies fall short of random assignment due to ethical or other concerns and utilize nonequivalent control groups (King, Vidourek, Davis, and McClellan 2002; Portwood et al., 2005; Slicker and Palmer, 1993). Others have not included a control group at all and/or have included only posttesting (Dappen and Isernhagen, 2006; Herrera, 2004; Terry, 1999).

One of the largest and most recent studies of school-based mentoring (SBM) that utilized an experimental design involved the Big Brother Big Sisters of America organization (BBBS) (Herrera, Grossman, Kauh, Feldman, McMaken and Jucovy, 2007). This rigorous national evaluation included 10 BBBS agencies, over 70 schools, and 1,139 youth in grades four through nine. It should be noted that all of the agencies selected to participate had a SBM program operating for at least four years, had strong leadership in place, and had well established relationships with the schools (best practices).

School staff typically referred the youth for participation in the SBM programs and many were identified as being economically or academically disadvantaged. Half of the students were randomly assigned to be matched with a volunteer mentor while the other half were placed on a waiting list to be matched after the study. The study involved a baseline and two follow-up surveys of students, their teachers, and the mentors. The researchers sought answers to several questions including “what benefits does BBBS SBM provide to youth socially, behaviorally, attitudinally, and academically?” and “what kinds of mentoring experiences help to ensure benefits?” The study measured impacts by comparing the progress of the youth in the treatment group to that of the control group youth.

In another recent study, Karcher (2008) examined the effects of adding school-based mentoring to other school-based support services already being provided in 19 schools. Students were randomly assigned to either receive a mentor in addition to other support services or the other supportive services alone. Participants were referred by parents, teachers, or themselves and received both pretesting and posttesting over the course of one school year. The study assessed outcomes identified in the literature as those most likely to be affected by school-based mentoring programs. These included math and reading grades, connectedness, self-esteem, social skills and support, hope, and how much youth feel they matter to others. This study does report that the agency involved did not provide many of the “best practices” that have been previously identified for achieving positive results (DuBois et al., 2002).

The U.S. Department of Education also published a large scale national impact evaluation of its Student Mentoring Programs in March 2009 (Bernstein, Dun Rappaport, Olsho, Hunt, and Levin, 2009). A team of independent contractual researchers conducted this experimental design study of 32 Student Mentoring Programs that included over 2,500 students in grades 4-8. Supported activities for programs funded under this competitive federal grant program are designed to improve interpersonal relationships, increase personal responsibility and community involvement, discourage the use of alcohol, drugs, and weapons, reduce drop-out rates, and improve academic performance. The evaluation report seeks to determine the impact of the programs in each of these areas on participants randomly assigned to receive services compared to a control group. Both self-report data and school records were collected at the beginning and end of the study school year. In addition to overall impacts, the study assesses impacts between subgroups of the participants.

Results from Prior Impact Evaluations

In their recent article, Rhodes and DuBois (2006) question whether the practice of mentoring has outpaced the research given the mixed results and documented implementation problems. In response they call for better alignment between research and practice and recommend policies that promote the use of evidence-based practices and rigorous evaluation. Findings from the above described evaluations do provide some encouragement for school-based mentoring programs. However, it is still unclear whether outcomes can last after services have ended.

These newer studies build upon previous work involving more established community-based mentoring (CBM) programs and early studies of school based programs. It is often cited that Tierney, Grossman, and Resch (1995) found strong evidence for a reduction in the use of alcohol and drugs, enhanced peer and child-parent relationships, better school attendance as well as improved attitudes about and performance in school.

A 2002 evaluation of the Healthy Kids Mentoring Program (King et al., 2002) reported significant improvements at posttest in mentored students' self-esteem levels and positive connections to school, peers, and family. In addition, mentored students had significantly higher school and family connectedness scores compared to nonmentored students at posttest. Portwood and her colleagues (2005) found evidence of improved school connectedness for participants in YouthFriends. This study also found improvements in community connectedness and goal setting for a subgroup of YouthFriends participants identified as at-risk.

The impact evaluation of Big Brothers Big Sisters school-based mentoring programs tested the extent to which school-based programs could provide youth with social, attitudinal, behavioral, and/or academic measurable benefits (Herrera et al., 2007). In general, the results were promising in that the study found positive outcomes for youth who participated in the program as measured by improved academic attitudes, performance, and behaviors. However, at the second follow-up many of the positive outcomes were not sustained.

At the end of the first school year teachers reported that the mentored students did better than the non-mentored group in several areas including overall academic performance, quality of work, number of assignments completed, and serious school infractions. The mentored youth also reported feeling more competent academically and less school skipping than their peers. Only these youth-reported outcomes held as significant at the second follow-up. The study does note that the differences between the two groups is small but comparable to results of the BBBS CBM study (Tierney et al., 1995). In addition, subgroup analyses were conducted but did not provide strong evidence for targeting services to specific groups of youth.

Karcher's (2008) impact evaluation revealed small positive effects for mentored students in terms of students' connectedness to peers, self-esteem, and social support from friends. However, he found no effect on academic outcomes. Subgroup analyses revealed that elementary boys and high

school girls benefited most from mentoring. Since this study examined mentoring as an additive service for students who already received other support services, the findings suggest that SBM “is of modest immediate value beyond other services provided to youth in schools and that it may have no direct, appreciable effect on academic achievement” (Karcher, 2008).

The impact study of the Student Mentoring Programs funded by the U.S. Department of Education found no statistically significant impacts on students for the sample as a whole on any of the three domains assessed: academic achievement and engagement, interpersonal relationships and personal responsibility, and high-risk or delinquent behavior (Bernstein et. al., 2009). However, additional subgroup analyses did reveal some positive outcomes for certain students. The program had a positive and significant impact on scholastic efficacy and school bonding, and pro-social behaviors for girls compared to boys. In addition, younger students in the mentoring group showed a significant improvement in terms of truancy compared to their nonmentored peers. The study also reports a low level of intensity in terms of service delivery for the Student Mentoring Programs.

In their 2010 report, Wheeler, Keller, and DuBois present a comparative analysis of the three recent large-scale school-based mentoring studies mentioned above. The authors suggest that school-based mentoring may be at a crossroads since arguments can be made both for and against further investment in these programs based on the interpretation of individual findings. Instead they aggregate results across the studies using meta-analytic techniques and show that there is some evidence of effectiveness on selected outcomes but not academic achievement. Their findings revealed positive outcomes on truancy, non-familial adult relationships, perceived scholastic efficacy, school-related misconduct, peer support, and absenteeism.

The HOPE CDC Mentoring Program

The HOPE Community Development Corporation (HOPE CDC) supports a faith-based initiative to improve

Figure 1. HOPE CDC’s Performance Measures

Objectives

- Improve academic performance, attendance, and instances of disciplinary referrals
- Improve interpersonal relationships
- Reduce the dropout rate
- Reduce juvenile delinquency and gang involvement

Measures

- Sustain student/mentor matches
- Improve student performance in core academic subjects
- Decrease unexcused absences from school
- Increase student GPAs
- Improve student attendance rates
- Reduce student disciplinary referrals
- Decrease suspensions and expulsions
- Help students refrain from drug use and violence

academic performance and behavior, reduce dropout rates, and generally prevent delinquency among at-risk youth. The in-school mentoring program evaluated here is only one aspect of the HOPE CDC’s overall efforts to prevent delinquency. While elements of both school-based and community-based mentoring programs can be found in the HOPE CDC’s model, the in-school mentoring program is primarily characterized as a school-based mentoring program.

In order to help these at-risk youth, HOPE CDC utilizes a variety of strategies involving teachers, parents, and other community resources. Mentors or Youth Development Specialists are assigned to the selected schools and are expected to provide a presence there. They first seek to develop positive relationships with the students by engaging them in various activities. Academic tutoring and lessons related to moral development and leadership skills are then incorporated into the meetings to encourage the youth to become better students and citizens. The curriculum underlying the HOPE mentoring program is TALKS (i.e.,

Transferring A Little Knowledge Systematically) (Davis, 2006). TALKS is designed to provide average adults with a method for effectively communicating with youth about respect, peer pressure, anger management, work ethics, and other relevant issues.

Performance indicators developed by the HOPE CDC highlight the goals they want to accomplish through the mentoring program. While the day-to-day operations of the program may vary by school, the overall goals of the program are the same. Figure 1 illustrates that the program established multiple objectives related to school performance and attendance as well as school behavior among youth. Regardless of the grade level for each school (i.e., elementary, middle, or high school), HOPE CDC aims to decrease unexcused absences, limit the number of disciplinary referrals, help students refrain from drug use and violence, and keep youth from being suspended or expelled from school. Moreover, an incentive or reward system is utilized by HOPE CDC to aid in the encouragement of students in these areas. By helping students with academic subjects and changing the attitudes of children and youth, HOPE CDC anticipates they can accomplish these program objectives. The data collected for this evaluation was designed to measure the program's impact around these domains.

As previously discussed, the process evaluation of the HOPE CDC's mentoring program indicates that implementation and integrity issues surrounding the program may hinder the ability to produce desired outcomes (see Haas and Turley, 2008). Specifically, the report concluded that the HOPE CDC program did not fully possess many of the elements found to be associated with successful mentoring programs. The model selected as a basis for the program was one concern of the evaluators as it was not found to be evidence-based and program staff departed from it in some substantial ways. Previous research confirms that it is those programs that exhibit "best practices" that are most successful at achieving positive results (DuBois et al., 2002; Herrera, Sipe, McClanahan, Arbeton, and Pepper, 2000; Grossman and Rhodes, 2002; Rhodes and DuBois, 2006; Jekielek, Moore, and Hair, 2002; MENTOR, 2003).

Methods

This report is the second study examining the quality of the HOPE CDC mentoring program. A previous process evaluation reported on the extent to which the HOPE CDC mentoring program engaged in practices shown to be important in the mentoring literature (see Haas and Turley, 2008). The results indicated that while there were positive aspects to the HOPE CDC's program it generally did not contain many of the characteristics shown to be empirically associated with successful mentoring programs. This evaluation examines the impact of the mentoring services on students' attendance, behavior, attitudes, and grades.

The HOPE CDC mentoring program operates in multiple cities and counties in West Virginia. The schools and student population that serve as the basis for this study, however, are all located in the city of Charleston. The mentoring program operates in six schools (i.e., one middle school, one high school, and 4 elementary schools) and works with students in the fourth, fifth, sixth, and ninth grades. This evaluation centers on the services provided by the HOPE CDC's mentoring program in these six schools during the 2007-2008 school year.

All four elementary schools have been identified by HOPE CDC as "Professional Development Schools." This status is largely determined by low scores among low income students on the WESTEST, a standardized achievement test for the state of West Virginia. According to county-level data obtained from the school district for the 2007-2008 school year, over 85.0% of students in these schools were identified as "needy" based on the percentage of students eligible for free or reduced lunch (West Virginia Department of Education, n.d.). In comparison, just over half of students were identified as "needy" in Kanawha County during the same year. The single high school in the evaluation had the smallest percentage of "needy" students at 49.8%. Nearly three-quarters of middle school students were defined as "needy" (74.1%).

In addition, the four elementary schools included in the study had a much greater minority and transient population compared to Kanawha County as a whole. During the 2007-

2008 school year, the elementary schools combined had a minority population varying between 60-70%, compared to the school population of the county at 15.0%. The middle and high schools were comprised of a 34.0% to 38.0% minority population (West Virginia Department of Education, n.d.).

To recruit students to participate in the HOPE CDC mentoring program and evaluation, school administrators were asked to identify students who they believed could benefit from mentoring services based on specific criteria. HOPE CDC requested that students be identified based on the following criteria: low grades, poor attendance, bad behavior, high number of disciplinary referrals, family issues, eligibility for free or reduced lunch, and low WESTEST scores. Once a list of students had been generated by the schools, two informed consent forms were sent to parents—one for enrollment in the program and one for enrollment in the evaluation study. A letter was also sent describing the HOPE CDC program, the procedures and data to be gathered as part of the evaluation, and their rights as a study participant. Upon receipt of consent forms from parents, the evaluation team worked with HOPE CDC staff to randomly assign students into treatment and control groups.

At the beginning of the 2007-2008 school year, all middle and high school students in the study group were asked to participate in a student survey. This information was collected prior to the initiation of mentoring services for the treatment group. In addition, school records pertaining to grades, attendance, and behavior were collected on all students for the 2006-2007 school year, the pretest period. This same information was collected again at the end of the study year for the posttest measurement.

Sample

A total of 129 students were ultimately enrolled in the study (i.e., 95 students in the fourth and fifth grades; 34 students in sixth and ninth grades). Students were then randomly assigned to either the treatment group or the control group. HOPE CDC began providing mentoring

Table 1. Study Group Demographics

Grade Level	Elementary Students				Middle/High School Students			
	Treatment Group		Control Group		Treatment Group		Control Group	
	n	%	n	%	n	%	n	%
4th Grade	25	52.1%	25	53.2%	----	----	----	----
5th Grade	23	47.9%	22	46.8%	----	----	----	----
6th Grade	----	----	----	----	6	33.3%	5	31.2%
9th Grade	----	----	----	----	12	66.7%	11	68.8%
Total	48	100.0%	47	100.0%	18	100.0%	16	100.0%
Gender								
Male	23	47.9%	20	42.6%	12	66.7%	8	50.0%
Female	25	52.1%	27	57.4%	6	33.3%	8	50.0%
Total	48	100.0%	47	100.0%	18	100.0%	16	100.0%
Race								
White	18	37.5%	10	21.3%	3	16.7%	7	43.8%
Nonwhite	30	62.5%	37	78.7%	15	83.3%	9	56.2%
Total	48	100.0%	47	100.0%	18	100.0%	16	100.0%
Age								
9	13	27.1%	15	31.9%	----	----	----	----
10	26	54.2%	19	40.4%	----	----	----	----
11	8	16.7%	11	23.4%	3	16.7%	1	6.2%
12	1	2.1%	2	4.3%	2	11.1%	3	18.8%
13	----	----	----	----	1	5.6%	1	6.2%
14	----	----	----	----	2	11.1	5	31.2%
15	----	----	----	----	8	44.4%	4	25.0%
16	----	----	----	----	2	11.1%	2	12.5%
Total	48	100.0%	47	100.0%	18	100.0%	16	100.0%
Free/Reduced Lunch								
Yes	10	20.8%	8	17.0%	7	38.9%	3	18.8%
No	38	79.2%	39	83.0%	11	61.1%	13	81.2%
Total	48	100.0%	47	100.0%	18	100.0%	16	100.0%
Times Held Back								
0	43	89.6%	42	89.4%	13	72.2%	12	75.0%
1	5	10.4%	5	10.6%	5	27.8%	4	25.0%
Total	48	100.0%	47	100.0%	18	100.0%	16	100.0%
Educational Placement								
Mainstream	38	79.2%	33	70.2%	10	55.6%	14	87.5%
Special	10	20.8%	14	29.8%	8	44.4%	2	12.5%
Total	48	100.0%	47	100.0%	18	100.0%	16	100.0%

Notes: Average age for both MS/HS treatment and control groups is 13.9. For Elementary students the average age is 9.9 for the treatment group and 10.0 for the control group. Only significant difference between the treatment and control groups is MS/HS educational placement ($p < .05$).

services to students in the treatment group while control group students were placed on a waiting list to receive services after the study period. Demographic information at pretest for each of the groups is provided in Table 1.

Students assigned to the treatment and control groups were statistically similar in terms of gender, race, and age at pretest. Slightly more of the elementary students were female in both groups. While not statistically significant, there were more males in the middle/high school treatment group than the control group. The majority of study participants in all groups were nonwhite. For elementary students the average age in the treatment group was 9.9 while control group students were an average of 10.0 years of age. The average age of middle/high school students in both groups was 13.9.

There were also no differences between the treatment and control groups for elementary or middle/high school students in terms of free/reduced lunch status or the number of times they had been held back. The majority of students were not receiving free/reduced lunch. Only about 10.0% of elementary students in either group had previously been held back a grade at pretest. About one-quarter of middle/high school students had been held back.

The only statistically significant difference between the treatment and control groups was on educational placement for the middle/high school students ($p=.041$). While the majority of students were in mainstream placement, 44.4% of those in the treatment group were receiving some type of special education. In the control group only 12.5% were receiving special services. No difference was indicated between the elementary treatment and control group students.

Data Sources

This report centers on the results of the impact evaluation. The research design and data collection methods applied in previous evaluations of mentoring programs helped to inform the approach used in the present study (e.g., Herrera et al., 2007; Karcher, 2008; Bernstein et al., 2009; Grossman and Rhodes, 2002; Herrera, 2004; Herrera et

al., 2000; Portwood et al., 2005; King et al., 2002; and Tierney et al., 1995). Hence, in accordance with previous evaluations, multiple data sources were used to obtain information about the students' progress after receiving mentoring services for roughly one school year. Data sources included official school records obtained from individual schools and the county board office as well as student surveys for middle and high school students.

The evaluation team cooperated with Kanawha County Schools to obtain data from the West Virginia Education Information System (WVEIS). Information on student grades, standardized test scores, attendance, behavior, and basic demographic characteristics were obtained from this system for the 2006-2007 and 2007-2008 school years. Data on school characteristics were also obtained from WVEIS.

In addition, information was also solicited from students. At the beginning and end of the 2007-2008 school year, students were asked to complete a survey. Only middle and high school students were asked to participate in the self-administered questionnaire. The student survey was completed in the classroom setting with research staff present. To complete the survey, students were called to a specific classroom identified by school staff. Student assent procedures were followed at that time. Of the 34 middle and high school students enrolled in the study, a total of 31 students participated in the pre-survey while 23 participated in the post-survey. Table 2 shows a breakdown of the demographic characteristics of the treatment and control group survey participants at pretest. No significant differences were found between the two groups. Responses provided personality, social learning, social bond, school and parental attachment, and perceptions of delinquency measures.

Measures

The student survey incorporates several different personal and social measures. A description of how the survey items were used to construct individual scales follows. Table 3 illustrates the Cronbach alpha reliability scores for each scale. All constructed scales exhibited a

Table 2. Student survey participants at Time 1

	Treatment Group		Control Group	
	n	%	n	%
Grade Level				
6th Grade	6	35.3%	4	28.6%
9th Grade	11	64.7%	10	71.4%
Total	17	100.0%	14	100.0%
Gender				
Male	11	64.7%	7	50.0%
Female	6	35.3%	7	50.0%
Total	17	100.0%	14	100.0%
Race				
White	2	11.8%	4	28.6%
Nonwhite	15	88.2%	10	71.4%
Total	17	100.0%	14	100.0%
Age				
11	3	17.6%	1	7.1%
12	3	17.6%	3	21.4%
13	0	0.0%	0	0.0%
14	3	17.6%	7	50.0%
15	6	35.3%	2	14.3%
16	2	11.8%	1	7.1%
Total	17	100.0%	14	100.0%
Grades				
Mostly A's/B's	3	18.8%	2	14.3%
Mostly B's/C's	4	25.0%	6	42.9%
Mostly C's/D's	7	43.8%	3	21.4%
Mostly D's/F's	1	6.2%	2	14.3%
Mostly F's	1	6.2%	1	7.1%
Total	16	100.0%	14	100.0%
Living Situation				
Mother & Father	3	17.6%	5	35.7%
Mother only	10	58.8%	2	14.3%
Father only	1	5.9%	2	14.3%
Father & stepmother	1	5.9%	0	0.0%
Grandparents	0	0.0%	3	21.4%
Other	2	11.8%	2	14.3%
Total	17	100.0%	14	100.0%
Home Ownership				
Yes	9	52.9%	7	50.0%
No	6	35.3%	6	42.9%
Don't know	2	11.8%	1	7.1%
Total	17	100.0%	14	100.0%

Notes: Official data used to fill in missing information on self-reported grade level, gender, race, and age. One participant in the treatment group did not report their grades. Average age of treatment group is 13.7, control group 13.6.

high level of reliability at both pre and posttest. An explanation of how official school records were used to measure attendance, behavior, and academic performance is also included.

Scales were constructed in six categories: personality, social learning, other, social bonds with school, social bonds with parents, and delinquency.

Personality

Impulsivity. This 4 item scale measures the degree to which students' actions are influenced by future goals. It asks whether students are more concerned with what happens in the short run or the long run and whether they tend to act on the spur of the moment. Items are measured on a 4-point Likert scale ranging from 1, strongly disagree to 4, strongly agree. Higher total scores indicate a greater degree of impulsivity.

Self-centeredness. This 4 item scale seeks to determine the students' level of sympathy towards the feelings and problems of others. Items ask whether the student looks out for himself and tries to get the things he wants regardless of the effect on other people. Items are measured on a 4-point Likert scale ranging from 1, strongly disagree to 4, strongly agree. Higher total scores indicate a greater degree of self-centeredness.

Risk seeking behavior. This 4 item scale measures the extent to which students' will engage in risky behavior just for fun or excitement. Students are asked whether excitement and adventure are more important than security. Items are measured on a 4-point Likert scale ranging from 1, strongly disagree to 4, strongly agree. Higher total scores indicate an increased level of risk seeking behavior.

Social Learning

Delinquent Peers. Here students were presented with 14 items describing delinquent behaviors and were asked to indicate how many of their closest friends, from 0 to 5, display those behaviors. For example, how many skip school without their parents permission, take things that don't belong to them, or get into physical fights. Responses were summed for the 14 items creating a scale ranging from 0 to

Table 3. Student Survey Reliability Analysis

	Items	Min	Max	Pretest		Posttest	
				n	alpha	n	alpha
<i>Personality</i>							
Impulsivity	4	4	16	26	.522	22	.696
Self-centeredness	4	4	16	28	.618	22	.625
Risk seeking behavior	4	4	16	28	.717	22	.651
<i>Social Learning</i>							
Delinquent Peers	14	0	70	30	.860	21	.927
<i>Other</i>							
Self efficacy	7	7	28	29	.641	23	.737
Social support	8	8	32	28	.869	22	.904
<i>Social Bonds - School</i>							
Attachment	8	8	48	30	.789	23	.840
Commitment	13	13	78	29	.852	22	.818
Beliefs	16	16	64	28	.744	22	.776
<i>Social Bonds - Parents</i>							
Trust	8	8	40	28	.909	22	.876
Alienation	5	5	25	27	.809	22	.744
Communication	4	4	20	28	.826	22	.797
Delinquency	16	0	64	29	.876	21	.847

70 where higher scores indicate greater delinquent peer associations.

Other

Self efficacy. This 7 item scale measures how confident the student is in dealing with unexpected events, solving problems, and generally handling whatever comes their way. Students are asked if they believe they have the coping skills to remain calm when faced with difficulties. Items are measured on a 4 point Likert scale ranging from 1, not true at all to 4, exactly true where higher scores indicate a greater degree of self efficacy.

Social support. This 8 item scale measures the degree to which students agree that they have people close to them who support and encourage them to do well. Students are asked if they have close family members, friends, and others

who can always be trusted, who understand their problems, and who help them feel good about themselves. Items are measured on a 4 point Likert scale ranging from 1, strongly disagree to 4, strongly agree.

Social bonds-School

Attachment to school. This 8 item scale measures the extent to which students feel as if they belong, look forward to going, and enjoy being in school. Students are also asked if they like their teachers and if they think homework is valuable. Items are measured on a 6 point Likert scale ranging from 1, strongly disagree to 6, strongly agree. Negatively worded statements were recoded so that higher total scores indicate a greater level of attachment.

Commitment to school. This 13 item scale measures whether students feel that it is important to work hard for

good grades, finish all of their homework and turn it in on time, and to graduate from school. Students are asked if getting good grades and other school activities are important to them. Items are measured on a 6 point Likert scale ranging from 1, strongly disagree to 6, strongly agree. Negatively worded statements were recoded so that higher total scores indicate a greater level of commitment.

Beliefs about school behavior. This 16 item scale measures if and how often students feel it is “OK” to break various school rules. For example, how often is it “OK” to be late for school, cheat on a test, talk back to teachers, or smoke on school grounds. Items are measured on a 4 point Likert scale ranging from 1, always to 4, never. Higher scores indicate that the student does not believe the behavior is acceptable at school and are therefore more positive.

Social bonds-Parents

Trust. This 8 item scale measures the degree to which students agree that they have an accepting respectful relationship with their parents/guardians. Students are asked if there is mutual trust and understanding between them and their parents. Items are measured on a 5 point Likert scale ranging from 1, strongly disagree to 5, strongly agree.

Alienation. This 5 item scale measures the students feelings toward their parents/guardians and whether they feel they get appropriate attention or credit from them. Students are asked if they are easily upset around their parents/guardians or if they get frustrated with their parents/guardians. All items were negatively worded and were thus recoded so that higher scores indicated lesser feelings of alienation. The 5 point Likert scale for these items then ranges from 1, strongly agree to 5, strongly disagree.

Communication. This 4 item scale measures whether students have open communication with their parents/guardians. Students are asked if they tell their parents about their problems and talk about their difficulties. Students are also asked if they can count on their parents when they need to get something off their chest. Items are measured on a 5 point Likert scale ranging from 1, strongly disagree to 5, strongly agree.

Self-reported Delinquency

This construct was measured by giving students a series of 16 statements about their own behaviors both in and out of school. Students were asked to indicate how frequently they engaged in each of the behaviors during the past 9 months. Behaviors ranged from cheating on a school test to using alcohol or tobacco to getting into trouble with the police. The response scale ranged from 0, not at all to 4, 10 or more times. Responses to all items were totaled to determine a delinquency score where higher values indicate more delinquent behavior.

Official School Records

Information for each of the students in the study was also obtained from official school records. These data included attendance, behavior, and academic performance records for the 2006-2007 (pretest) and 2007-2008 (posttest) school years.

From student attendance records obtained from the Kanawha County Board Office, a simple count of the number of tardies, excused absences, and unexcused absences was determined for each student for each time period.

A total behavior score based on the Kanawha County Schools Respect and Protect Policy was computed for each student in the study. Reported violation codes were obtained for the pretest and posttest school years. For the elementary students there were between 0 and 7 recorded offenses per student during the pretest school year. At posttest the number of recorded violations increased to between 0 and 77. Middle/high school students had between 0 and 15 recorded violations at pretest. At posttest the maximum number of violations recorded for this group also increased to 30.

The violation codes can be categorized into four levels where level 1 is the least serious and level 4 is the most serious. Disruptive behavior, disobeying class rules, and cheating are examples of level 1 violations. Level 4 violations could involve possession, use, or distribution of illegal drugs/substances, battery against a school employee, or possession of a weapon. Violation codes were recoded in the data to the numeric value of the level of the violation. For those codes that could be categorized in more than one

level, the least serious value was assigned. For example, the code “BDT” could be a level 2 violation, disobeying a school staff member in a willful manner. It is also included in the level 3 violations as *habitually* disobeying a school staff member in a willful manner. For consistency, the code “BDT” was always assigned to level 2 since it was not easily determined if the offense should be considered habitual. The values were then summed so that the behavior score accounts for both frequency and severity of offenses for each student.

Academic performance was measured in terms of scores on the standardized achievement test for the state, WESTEST, and letter grades for math and reading/English. The research team attempted to obtain both percentage and letter grades for these subjects from individual schools as well as the county board office; however, percentages were not available for the majority of students. Also, letter grades could not be obtained for about half of the middle/high school students.

WESTEST scores for math, reading, science, and social studies were obtained and compared across groups and over time. In addition, WESTEST scores are discussed in relation to the five established performance levels based on student competency ranging from novice to distinguished. The high school, or ninth grade, students in the study did not take the WESTEST during the 2007-2008 school year. Thus, in order to obtain a posttest measure for these students, their results from the 2008-2009 school year had to be used. However, an updated version of the test, the WESTEST 2, was administered to all students in that year. Because of this difference, additional analyses were conducted to examine the middle and high school results separately.

Each student's letter grade in math and reading/English was converted to the standard 0-4 point scale where an “F” (sometimes indicated as “E”) equals 0 and an “A” equals 4. The two values were then combined to create a total letter grade score at pretest and posttest.

Results

In order to assess the effectiveness of the HOPE School-Based Mentoring program, the posttest scores of students in the treatment group were compared to those of the control group on several different constructs. A comparison of the two groups at pretest indicated some differences. In addition, the mean difference over time within each group was also examined. This method of analysis was used for both self-reported student survey data and official school data. Overall little if any significant difference was observed between the treatment and control groups after implementation of the mentoring initiative.

Self-reported Student Survey Data

No significant differences were observed between the treatment and control groups at posttest on any of the social bond measures (Table 4). Students who participated in the mentoring program were no more likely to have strong bonds to school or parents than students in the control group. However, mean scores were fairly high for both groups indicating strong attachments and commitments to school as well as strong beliefs about acceptable behavior at school. This is true at both pretest and posttest. The only significant difference observed between the two groups was on beliefs about acceptable school behavior at pretest. Here the difference in mean scores favored the control group ($t = 2.29, p = .031$).

Likewise, all students indicated high levels of trust and communication with their parents/guardians. High mean scores on alienation indicate that students did not feel alienated from their parents/guardians. While the differences between the two groups at posttest were not significant, mean scores for the treatment group increased for both trust and alienation. Mean scores for the control group declined slightly on each of these measures from pre to posttest. Communication scores were only slightly lower at posttest for both treatment and control group students.

Students in both the treatment and control groups reported little involvement in school activities. When asked whether they participated in 11 different types of school

Table 4. Social Bonds

	Pretest		Posttest		df	t
	Mentor	Control	Mentor	Control		
School						
Attachment	32.94 (7.01) <i>n</i> = 16	36.86 (6.67) <i>n</i> = 14	34.36 (10.31) <i>n</i> = 11	35.42 (5.33) <i>n</i> = 12	21	.312
Commitment	60.00 (7.77) <i>n</i> = 16	59.69 (11.76) <i>n</i> = 13	59.00 (10.69) <i>n</i> = 11	59.73 (4.05) <i>n</i> = 11	20	.211
Beliefs	58.07 (4.06) <i>n</i> = 15	61.08 (2.63) <i>n</i> = 13	59.09 (4.01) <i>n</i> = 11	58.73 (2.80) <i>n</i> = 11	20	-.247
Involvement	2.88 (2.03) <i>n</i> = 17	2.79 (1.81) <i>n</i> = 14	2.73 (2.90) <i>n</i> = 11	3.33 (1.72) <i>n</i> = 12	21	.616
Relationship w/ teachers	5.75 (1.88) <i>n</i> = 16	7.69 (1.60) <i>n</i> = 13	4.64 (3.26) <i>n</i> = 11	6.83 (1.90) <i>n</i> = 12	15.79	1.95
Parents						
Trust	32.86 (7.80) <i>n</i> = 14	32.86 (5.90) <i>n</i> = 14	35.00 (4.96) <i>n</i> = 11	31.45 (6.12) <i>n</i> = 11	20	-1.493
Alienation	18.15 (5.05) <i>n</i> = 13	18.71 (4.01) <i>n</i> = 14	19.73 (4.10) <i>n</i> = 11	18.09 (3.33) <i>n</i> = 11	20	-1.027
Communication	14.64 (4.34) <i>n</i> = 14	14.57 (3.98) <i>n</i> = 14	14.18 (4.85) <i>n</i> = 11	13.45 (3.14) <i>n</i> = 11	20	-.417
Relationship w/ parents	8.50 (2.13) <i>n</i> = 16	8.57 (1.60) <i>n</i> = 14	8.82 (1.66) <i>n</i> = 11	8.08 (1.62) <i>n</i> = 12	21	-1.073

Note:
Shown mean (standard deviation) and number of respondents.

activities, students in both groups reported involvement in less than 3 activities at pretest. Mean scores were basically the same at posttest slightly favoring the control group (3.33) over the treatment group (2.73).

At pretest the control group students (7.69) rated their relationship with their teachers significantly better than the treatment group students (5.75). However, the difference between the two groups was not significant at posttest and still favored the control group. Scores for both groups did fall in the middle to high range of the scale at both measurement times.

Students in both groups rated their relationships with their parents/guardians positively. Mean scores for students in the treatment group increased slightly from pretest (8.50)

to posttest (8.82). On the other hand, students in control group had a slightly lower mean score at posttest (8.08) than at pretest (8.57).

Table 5 further examines personal and social characteristics of the two study groups at pre and posttest. Again, no significant differences were found between the treatment and control groups at posttest. The treatment group (10.13) was; however, significantly more likely to exhibit self-centeredness than the control group (8.23) at pretest ($p = .039$). At posttest the mean score for the treatment group fell to 6.90 on this measure while the control group score declined only slightly to 8.08.

On the other two personality measures, impulsivity and risk taking behavior, the treatment group improved somewhat from pretest to posttest. Conversely, mean scores for the

control group increased slightly indicating greater impulsivity and more risk taking behavior. This resulted in a lower mean score for the treatment group (9.20) than the control group (9.42) on impulsivity at posttest. In terms of risk taking behavior, mean scores favored the treatment group at both pretest and posttest. There was, however, no significant difference between the two groups at posttest on either measure. Scores for both groups were at or just above the middle of the scale for these characteristics.

Students in the treatment and control groups were nearly evenly matched in terms of self efficacy and social support. Both groups indicated high levels of social support with an average score of about 27 at pretest (out of a maximum of 32). Results at posttest slightly favored the control group (28.67) over the treatment group (26.90). Likewise, self-efficacy scores were on the high end of the scale for both the treatment (20.13) and control group (21.43) at pretest. There was a slight decline at posttest for the treatment group (18.64) while the control group reported a somewhat higher level of self efficacy (21.75).

Generally, students participating in the study did not characterize their peers as participants in delinquent behavior. Social learning/delinquent peers mean scores were low for both treatment (17.31) and control group (8.86) students at pretest. This is, however, a significant difference between the two groups ($p = .043$). At posttest the mean for the control group increased to 20.55 while the treatment group increased only slightly to 20.10. While the difference favors the treatment group at posttest it is not significant.

In terms of their own delinquency, students reported little or no participation in delinquent behavior. The difference between the two groups was, however, significant at pretest but not at posttest. As shown in Graph 1, students in the treatment group had a mean delinquency score of 8.56 at pretest while control group students averaged just 3.00. After participation in the mentoring program, the average delinquency score for students in the treatment group declined to 5.40. Control group students reported a higher level of delinquency (3.73) at posttest but still registered extremely low on the scale.

Official School Data

Next this study examined official records of student attendance, behavior, and academic performance both prior to and after the mentoring program. For these analyses, elementary students are discussed separately from the middle/high school group.

While no significant differences in attendance were noted between the elementary treatment and control groups at pretest or posttest, there were differences within each group over time. At pretest, students assigned to the treatment group had more tardies, excused absences, and unexcused absences than their counterparts in the control group (Table 6). At posttest, treatment group students still had more reported absences (both excused and unexcused). However, fewer tardies were reported for the treatment group (8.74) than the control group (8.98) at posttest.

For both elementary groups, the mean difference in tardies improved and was significant from pretest to posttest. For students in the treatment group, the mean difference

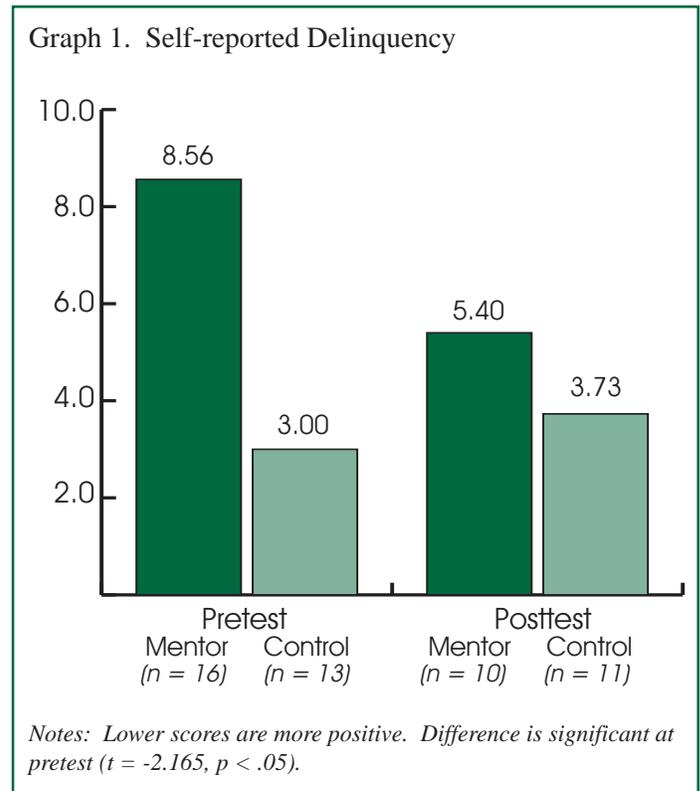


Table 5. Personality and Social Learning

	Pretest		Posttest		df	t
	Mentor	Control	Mentor	Control		
<i>Personality</i>						
Impulsivity	10.47 (2.48) <i>n</i> = 15	8.82 (2.23) <i>n</i> = 11	9.20 (3.55) <i>n</i> = 10	9.42 (2.35) <i>n</i> = 12	20	.171
Self-centeredness	10.13 (2.30) <i>n</i> = 15	8.23 (2.32) <i>n</i> = 13	6.90 (1.79) <i>n</i> = 10	8.08 (2.68) <i>n</i> = 12	20	1.19
Risk taking behavior	8.80 (3.01) <i>n</i> = 15	9.38 (2.18) <i>n</i> = 13	7.80 (2.30) <i>n</i> = 10	9.75 (3.17) <i>n</i> = 12	20	1.621
<i>Social Learning</i>						
Delinquent Peers*	17.31 (13.12) <i>n</i> = 16	8.86 (9.74) <i>n</i> = 14	20.10 (13.86) <i>n</i> = 10	20.55 (21.19) <i>n</i> = 11	17.36	.057
<i>Other</i>						
Self efficacy	20.13 (4.19) <i>n</i> = 15	21.43 (2.56) <i>n</i> = 14	18.64 (4.82) <i>n</i> = 11	21.75 (2.26) <i>n</i> = 12	21	2.011
Social support	27.20 (5.27) <i>n</i> = 15	27.62 (3.69) <i>n</i> = 13	26.90 (5.24) <i>n</i> = 10	28.67 (3.85) <i>n</i> = 12	20	.912

Note: Shown mean (standard deviation) and number of respondents. *Lower scores on delinquent peers are more positive.

illustrates 7.64 fewer tardies at posttest than at pretest ($p = .002$). Control group students also showed improvement at posttest with 3.35 fewer tardies than at pretest ($p = .026$). Thus, the change over time in tardies was more favorable for the treatment group.

Similar results were observed for unexcused absences in that both elementary groups reported fewer at posttest than pretest. For students in the treatment group, however, the mean difference over time was significant and greater than that for the control group. The mean difference for treatment group students illustrates 2.36 fewer unexcused absences at posttest ($p = .028$). For the control group the mean difference over time was just 0.85.

Little change in reported excused absences was observed over time for elementary students in either the treatment or control group. While not significant the mean number of excused absences increased from 4.65 to 4.98 for the treatment group and from 3.68 to 3.87 for the control group.

This result is not unexpected since excused absences are more likely to be for legitimate reasons. It wouldn't seem that these should differ significantly over time.

Attendance results for middle/high school students also exhibited no significant differences between the treatment and control groups at either pretest or posttest. All three measures, tardies, excused absences, and unexcused absences, favored the control group at pretest. That is there were fewer reported tardies, excused absences, and unexcused absences, for students in the control group compared to those in the treatment group at pretest. This held true at posttest with the exception of unexcused absences. While the mean number of unexcused absences increased rather substantially at posttest for both groups, the treatment group (36.00) had fewer reported on average than the control group (38.17).

The difference in reported unexcused absences over time was significant for both the treatment and control groups.

Table 6. Attendance and Behavior

	Pretest		Posttest		df	t
	Mentor	Control	Mentor	Control		
Elementary						
Tardies	16.62 (19.88) <i>n</i> = 48	12.30 (15.97) <i>n</i> = 47	8.74 (13.09) <i>n</i> = 47	8.98 (10.92) <i>n</i> = 46	91	.093
Excused Absences	4.65 (8.12) <i>n</i> = 48	3.68 (4.44) <i>n</i> = 47	4.98 (6.98) <i>n</i> = 47	3.87 (4.77) <i>n</i> = 46	91	-.893
Unexcused Absences	9.48 (10.15) <i>n</i> = 48	5.87 (8.29) <i>n</i> = 47	6.60 (6.95) <i>n</i> = 47	4.80 (5.76) <i>n</i> = 46	91	-1.352
Behavior	1.90 (4.12) <i>n</i> = 48	1.87 (3.75) <i>n</i> = 47	6.29 (16.83) <i>n</i> = 48	6.09 (17.20) <i>n</i> = 47	93	-.059
Middle/High School						
Tardies	5.12 (7.76) <i>n</i> = 17	3.00 (2.90) <i>n</i> = 16	6.33 (6.31) <i>n</i> = 15	5.00 (6.30) <i>n</i> = 15	25	-.546
Excused Absences	7.06 (11.07) <i>n</i> = 17	4.13 (4.11) <i>n</i> = 16	10.40 (16.54) <i>n</i> = 15	6.75 (8.63) <i>n</i> = 12	25	-.691
Unexcused Absences	13.65 (14.13) <i>n</i> = 17	11.94 (13.25) <i>n</i> = 16	36.00 (21.59) <i>n</i> = 15	38.17 (32.96) <i>n</i> = 12	25	.206
Behavior	7.53 (6.87) <i>n</i> = 17	4.38 (7.94) <i>n</i> = 16	12.41 (8.65) <i>n</i> = 17	9.19 (11.74) <i>n</i> = 16	31	-.902

Notes: Shown mean (standard deviation) and n. Independent samples t test shown for posttest only. No significant differences between groups at posttest.

However, the difference was negative for both groups meaning that they had a greater number of unexcused absences at posttest than at pretest. The mean difference for the treatment group indicates on average 21.0 more unexcused absences at posttest. For the control group the mean difference shows an increase of on average 27.83 unexcused absences.

Behavior scores are based on recorded disciplinary violations under the schools' respect and protect policies and account for both frequency and severity. Generally, few violations were reported for both elementary and middle/

high school students participating in this study. However, there were increases across all groups from pretest to posttest.

Elementary students in both the treatment and control groups had an average behavior score of less than 2.0 during the school year preceding the study. At posttest the treatment group had increased to an average score of 6.29 while the control group increased to 6.09. The difference between the two groups was not significant at either time.

The difference over time was also not significant within either group of elementary students. Both groups fared worse at posttest and the mean difference in scores was roughly the same. For the treatment group the mean behavior score

increased on average by 4.40. The increase over time for the control group was 4.21.

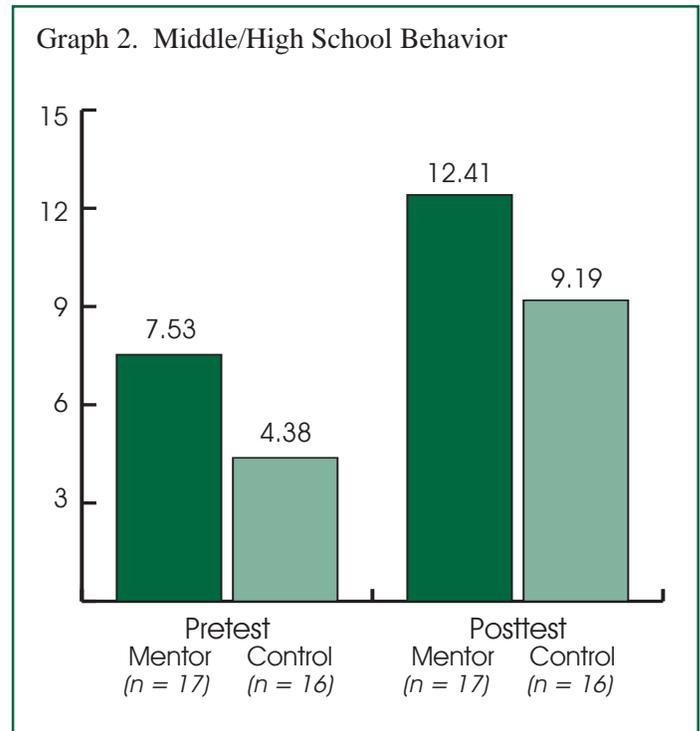
Middle/high school students were also worse in terms of behavior at posttest than at pretest. Mean behavior scores favored the control group at both measurement times. At pretest a mean behavior score of 7.53 was reported for the treatment group while the control group averaged 4.38. Both had increased at posttest with the treatment group (12.41) still having a more negative score than the control group (9.19). The difference between the two groups was not significant at either measurement time.

The mean difference over time was not significant for either the treatment or the control group of middle/high school students. Behavior scores also increased at posttest for both groups with nearly the same mean difference (4.8).

Finally, the treatment and control groups were compared at pretest and posttest in terms of academic performance by examining WESTEST scores and grades. Table 7 illustrates WESTEST scores in the four core subject areas for elementary and middle/high school students at pre and posttest. In addition, when available letter grades for math and reading (or English) were combined into a numeric score for the students' grade in these subjects.

Elementary students in both the treatment and control groups did better across the board at posttest on the WESTEST. However, there were no significant differences between the two groups at either pre or posttest. Scores favored the treatment group slightly in all four subject areas at pretest. This held true at posttest with the exception of science scores where the groups scored fairly even.

Cut score ranges are established for the WESTEST to denote which of the five performance levels (from novice to distinguished) the student falls into. These ranges are based on content area and grade level with the middle range defined as mastery, meaning the student shows competent performance. For the purposes of this study, scores for both fourth and fifth grade students are combined in the elementary posttest results shown. For comparison it can, however, be said that the average scores at posttest for students in both the treatment and control groups were



generally within the mastery ranges for each content area.

The change over time was significant across all subject areas for both groups of elementary students. Given that the control group students started out with slightly lower scores at pretest, mean differences mostly favored this group. In math, science, and social studies the mean difference from pre to posttest was greater for the control group. Control group students raised their scores by 32.48 points in math, 21.58 in science, and 21.16 in social studies. This compares to increases of 31.84 points in math for the treatment group and 17.52 and 16.86 in science and social studies respectively. Treatment group students had the greater mean difference in reading scores across time. Here the treatment group improved their scores by 24.98 points compared to 21.26 for the control group.

Letter grades in math and reading for elementary students also showed that the treatment group did somewhat (though not significantly) better than the control group at both pre and posttest. Grades for both declined from pretest to posttest. At pretest the math/reading grade score was 4.58 for students in the treatment group compared to 4.33 for

Table 7. Academic Performance

	Pretest		Posttest		df	t
	Mentor	Control	Mentor	Control		
Elementary						
<i>WESTEST Scores</i>						
Math	618.85 (28.68) <i>n</i> = 46	611.39 (33.67) <i>n</i> = 46	652.50 (28.87) <i>n</i> = 46	643.39 (31.19) <i>n</i> = 44	88	-1.439
Reading	613.59 (39.15) <i>n</i> = 46	610.74 (42.58) <i>n</i> = 46	641.00 (22.94) <i>n</i> = 46	631.34 (27.04) <i>n</i> = 44	88	-1.830
Science	610.28 (22.41) <i>n</i> = 46	607.43 (29.87) <i>n</i> = 46	629.30 (29.12) <i>n</i> = 46	629.68 (27.11) <i>n</i> = 44	88	.064
Social Studies	614.27 (40.30) <i>n</i> = 45	613.20 (31.01) <i>n</i> = 46	635.11 (26.09) <i>n</i> = 46	633.82 (20.85) <i>n</i> = 44	88	-.259
Math/Reading grade	4.58 (2.02) <i>n</i> = 48	4.33 (1.64) <i>n</i> = 45	4.28 (1.74) <i>n</i> = 47	3.85 (1.81) <i>n</i> = 47	92	-1.163
Middle/High School						
<i>WESTEST Scores</i>						
Math	637.24 (27.67) <i>n</i> = 17	675.27 (39.07) <i>n</i> = 15	600.80 (51.83) <i>n</i> = 15	617.36 (56.11) <i>n</i> = 11	24	.778
Reading	633.59 (30.13) <i>n</i> = 17	666.07 (24.67) <i>n</i> = 14	438.27 (147.63) <i>n</i> = 15	501.78 (161.31) <i>n</i> = 9	22	.986
Science	642.14 (37.63) <i>n</i> = 14	663.87 (44.09) <i>n</i> = 15	569.00 (77.42) <i>n</i> = 14	603.20 (60.29) <i>n</i> = 10	22	1.165
Social Studies	639.43 (35.46) <i>n</i> = 14	669.00 (26.19) <i>n</i> = 15	464.14 (144.63) <i>n</i> = 14	474.50 (120.96) <i>n</i> = 10	22	.185
Math/English grade	2.61 (1.91) <i>n</i> = 18	3.80 (2.34) <i>n</i> = 15	1.71 (1.70) <i>n</i> = 7	2.44 (2.40) <i>n</i> = 9	14	.680

Notes: Letter grades were converted to a 4-point scale and a combined score for math and reading was calculated. Unable to obtain letter grades for over half of the middle/high school students at posttest. No significant differences at posttest.

students in the control group. The score for students in the treatment group dropped to 4.28 at posttest. The decline for the control group to 3.85 at posttest was significant.

WESTEST scores for the middle/high school students generally favored the control group at both pretest and posttest. The difference between the treatment and control groups was also statistically significant at pretest for math ($p = .003$), reading ($p = .003$), and social studies ($p = .016$). While not significant, the control did better than the treatment group in all four subject areas at posttest.

It is difficult to make comparisons across time for the middle/high school group due to data collection constraints. For the high school students in this group, the WESTEST scores are an additional year after the mentoring program took place and involve a different version of the test. Thus, while it may appear that students scored much worse at posttest, this may not be the case. Established cut score ranges for the WESTEST 2 are lower for the high school students, particularly for reading and social studies.

When the WESTEST scores are examined separately for the middle and high school students, the overall differences are explained mainly by the high school group. Accounting for the differences in cut score scales, both the treatment and control groups of high school students had lower performance levels for math and reading at posttest. However, the control group declined from mastery to partial mastery while the treatment group fell from partial mastery to novice at posttest. The treatment group also fell to novice in social studies while the control group remained at mastery. Science cut scores were not available for this grade level on the WESTEST 2.

Little variation was observed for the middle school students either between groups or over time. Middle school students in the control group declined from mastery to partial mastery in math and social studies at posttest but increased from partial mastery to mastery in science. Reading scores for this group were at the mastery level at both pre and posttest. The middle school students in the treatment group were at partial mastery for all subjects at pretest and only increased to mastery in science at posttest.

Letter grade scores for middle/high school students again favored the control group at both pre and posttest; however, the differences between groups was not significant. At pretest, the math/English grade score for students in the control group was 3.80 compared to 2.61 for treatment group students. By posttest the control group score had declined to 2.44 while the score for students in the treatment group fell to 1.71.

It should be noted that letter grades were not provided for all middle/high school students at posttest. Thus caution should be used when making comparisons across time since only about half of the students are included in the posttest mean results. However, once this is accounted for, the mean difference from pretest to posttest is the same (1.00) for both the treatment and control groups and is not significant.

Discussion and Conclusions

This evaluation examines a very new program. Even though program administrators had previously established relationships with some of the schools, this specific program was in its first year. It is, therefore, not surprising that many of the implementation issues common to new prevention and intervention programs were discovered through this evaluation.

Previously, the results of the process evaluation indicated that few of the known “best practices” were present in the HOPE program. This is of concern because previous research on mentor programs has highlighted the importance program design and implementation for achieving the desired outcomes. In fact, even the best designed and implemented school-based mentoring programs have produced relatively small changes in youth school performance and behavior. This further underscores the importance of program quality for achieving behavior change among mentees. Given the timing of this evaluation as well as the programmatic issues identified through the process evaluation, it was questionable whether the HOPE was operating in a manner that would result in positive behavior changes among youth.

Prior research informed the research design and methodology utilized in the present study. Previous researchers and experts on mentor programs have called for the use of rigorous methods in conducting impact evaluations. This study replicated many of the same methods used in various large-scale national evaluations of mentor programs. Many of the best designed studies utilize random assignment for treatment and control groups, some form of a waiting list for nonselected participants, and included a combination of official school records and self-report data from students.

Generally, posttest analysis yielded little or no differences in school performance and behavior measures between students who participated in the HOPE program and students in the control group. However, this may have been due in part to differences in the two groups at pretest. At pretest, the control group reported more positive beliefs about school behavior and relationships with their teachers, less self-centered tendencies, and fewer delinquent acts for both themselves and their peers.

Relatively high scores at pretest on many of the outcome measures may have also limited the capacity of the study to identify significant changes in across groups. As noted previously, many of the pretest scores were high on various measures making it more difficult to ascertain impact over time. The high scores at pretest generally indicated that both the treatment and control group were doing fairly well on the constructs prior to the intervention. For instance, high levels of attachment and commitment to school, trust and communication with parents, relationships with teachers and parents, and social support were exhibited by both groups at pretest. Students also reported little or no delinquent behavior for themselves or their friends. Such behaviors remained low at posttest as well. Both groups reported little involvement in school activities at both pre and posttest.

The analysis of official school data yielded a similar story as well. Little or no differences were found at posttest between the two groups and when differences were present they tended to favor the control group. The number of tardies decreased significantly for both groups of elementary students

at posttest. Similarly, the number of unexcused absences decreased at posttest for both elementary groups; however, the decline was significant only for the treatment group. The difference in unexcused absences over time was also significant for both treatment and control group students in middle/high school. However, the difference was negative for both groups indicating a significant increase in unexcused absences at posttest. The number of tardies was also slightly greater at posttest for both groups of middle/high school students.

Behavior scores were generally low for all students in the study indicating few disciplinary violations. Surprisingly, however, there were generally more violations reported at posttest than at pretest. None of the findings regarding student behavior were significant either between the two groups or within the groups over time. The mean difference over time within both groups was negative (i.e., indicating more behavior infractions) and roughly the same for both elementary and middle/high school students.

In terms of academics, WESTEST scores for elementary students in both the treatment and control groups improved significantly at posttest. There was, however, no difference between the groups and it should be noted that all students were doing well with scores generally falling in the mastery level. The difference between groups of middle/high school students was significant at pretest and favored the control group. Problems with data collection make it difficult to determine posttest differences for middle/high school students.

While it is not clear that the HOPE program had a significant impact on the school performance and behavior of students in this study, it is hoped that the results of this evaluation will provide valuable information for improving the program. The results of this study, coupled with what was learned through the prior process evaluation, should provide program and school administrators with a wealth of information on the practices which constitute the most effective school-based programs and highlight much needed areas of improvement. What is known is that many children and youth need strong schools and the presence of active community leaders in their lives. It is hoped that through

continued program development and evaluation, the HOPE program will modify practices to be more in-line with programs that have yielded success in the past and as a result better meet the needs of at-risk youth in the future.

References

- Bernstein, L., Dun Rappaport, C., Olsho, L., Hunt, D., and Levin, M. (2009). *Impact evaluation of the U.S. Department of Education's Student Mentoring Program*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Dappen, L., and Isernhagen, J.C. (2006). Urban and nonurban schools examination of a statewide student mentoring program. *Urban Education, 41* (2), 151-168.
- Davis, H. D. (2006). TALKS mentoring movement-TALKS mentoring of Champaign County-Transferring a little knowledge systematically. Champaign, IL: KJAC Publishing.
- DuBois, D. L., Holloway, B. E., Valentine, J. C., and Cooper, H. (2002). Effectiveness of mentoring programs for youth: A meta-analytic review. *American Journal of Community Psychology, 30*, 157-197.
- Grossman, J. B., and Rhodes, J. E. (2002). The test of time: Predictors and effects of duration in youth mentoring relationships. *American Journal of Community Psychology, 30* (2), 199-219.
- Haas, S.M., and Turley, E. (2008). *Helping others pursue excellence: A process evaluation of HOPE CDC's mentoring program*. Charleston, WV: Criminal Justice Statistical Analysis Center, Division of Criminal Justice Services, Department of Military Affairs and Public Safety.
- Herrera, C. (2004). *School-based mentoring: A closer look*. Philadelphia, PA: Public/Private Ventures.
- Herrera, C., Grossman, J. B., Kauh, T. J., Feldman, A. F., McMaken, J., and Jucovy, L. Z. (2007). *Making a difference in schools: The Big Brothers Big Sisters school-based mentoring impact study*. Philadelphia, PA: Public/Private Ventures.
- Herrera, C., Sipe, C. L., McClanahan, W. S., Arbeton, A. J., and Pepper, S. K. (2000). *Mentoring school-age children: Relationship development in community-based and school-based programs*. Philadelphia, PA: Public/Private Ventures.
- Jekielek, S., Moore, K. A., and Hair, E. C. (2002). *Mentoring programs and youth development: A synthesis*. Washington, DC: Child Trends.
- Karcher, M. J. (2008). The study of mentoring in the learning environment (SMILE): A randomized evaluation of the effectiveness of school-based mentoring. *Prevention Science, 9*, 99-113.
- Karcher, M. J., Kuperminc, G. P., Portwood, S. G., and Sipe, C. L. (2006). Mentoring programs: A framework to inform program development, research, and evaluation. *Journal of Community Psychology, 34* (6), 709-725.
- King, K. A., Vidourek, R. A., Davis, B. and McClellan W. (2002). Increasing self-esteem and school connectedness through a multidimensional mentoring program. *Journal of School Health, 72* (7), 294-299.
- MENTOR/National Mentoring Partnership. (2003). *Elements of Effective Practice (2nd Edition)*. Alexandria, VA. Retrieved September 12, 2008, from www.mentoring.org.
- Portwood, S. G., Ayers, P. M., Kinnison, K. E., Waris, R. G., and Wise, D. L. (2005). YouthFriends: Outcomes from a school-based mentoring program. *The Journal of Primary Prevention, 26* (2), 129-145.
- Rhodes, J. E. and DuBois, D. L. (2006). Understanding and facilitating the youth mentoring movement. *Society for Research in Child Development Social Policy Report, 20* (3), 1-19.

Slicker, E. K., and Palmer, D. J. (1993). Mentoring at-risk high school students: Evaluation of a school-based program. *School Counselor*, 40 (5), 327-334.

Terry, J. (1999). A community/school mentoring program for elementary students. *Professional School Counseling*, 2 (3), 237-240.

Tierney, J. P., Grossman, J. B., and Resch, N. L. (1995). *Making a difference: An impact study of Big Brothers Big Sisters*. Philadelphia, PA: Public/Private Ventures.

Wheeler, M. E., Keller, T. E., and DuBois, D. L. (2010). Review of three recent randomized trials of school-based mentoring: Making sense of mixed findings. *Society for Research in Child Development Social Policy Report*, 24 (3), 3-21.

West Virginia Department of Education (n.d.). West Virginia Education Information System (WVEIS) Data On-Demand. Accessed online at: <http://wveis.k12.wv.us/nclb/pub/index.cfm>.

Acknowledgments

The authors would like to thank Rev. Matthew Watts, *HOPE CDC President*, Bonita Perry-Dean, *HOPE CDC Chief of Staff* and all of the HOPE CDC Youth Development Specialists and Coordinators for their support and assistance with the program documentation review and interview process.

We would also like to thank Dr. Ronald Duerring, *Superintendent of Kanawha County Schools* for allowing the research staff access to schools and assistance in obtaining WVEIS data. James Withrow, Jerry Legg, and Nancy Baldwin, of the Kanawha County Board of Education provided and assisted with the interpretation of the WVEIS data.

School administrators and staff at Chandler Elementary, Glenwood Elementary, J.E. Robins Elementary, Piedmont Elementary, Stonewall Jackson Middle School, and Capital High School are greatly appreciated for allowing us access to their schools and for their participation and cooperation throughout the interview and survey process.

Finally, the authors would like to acknowledge Monika Sterling, *Statistical Analysis Center* for her invaluable assistance with the many data collection procedures involved with this project.

Funding Source

This project was supported by grant #2007-BJ-CX-K052, awarded by the U.S. Department of Justice, Bureau of Justice Statistics through the State Justice Statistics Program. The views expressed in this report are those of the authors and do not necessarily reflect the opinions of the U.S. Department of Justice, the Bureau of Justice Statistics, or the West Virginia Division of Justice and Community Services.

Recommended Citation

Haas, S. M. and Turley, E., (2011). *Helping Others Pursue Excellence in Public Schools: Assessing the Impact of HOPE CDC's Mentoring Program*. Charleston, WV: Criminal Justice Statistical Analysis Center, Office of Research and Strategic Planning, Division of Justice and Community Services, Department of Military Affairs and Public Safety. Available online at www.djcs.wv.gov/SAC.



DJCS Administration

J. Norbert Federspiel, *Director*
Jeff Estep, *Chief Deputy Director*
Stephen M. Haas, *Deputy Director*
Leslie Boggess, *Deputy Director*

ORSP Administration

Stephen M. Haas, *Director*

1204 Kanawha Boulevard, East
Charleston, WV 25301
Phone: 304-558-8814
Fax: 304-558-0391
www.djcs.wv.gov