# Impact Evaluation of the U.S. Department of Education's Student Mentoring Program

## Final Report

![ies logo] NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE

Institute of Education Sciences

# Impact Evaluation of the U.S. Department of Education's Student Mentoring Program

Final Report

**MARCH 2009**

**Lawrence Bernstein, Project Director**
**Catherine Dun Rappaport**
**Lauren Olsho**
**Dana Hunt**
**Marjorie Levin**
Abt Associates

With the assistance of:
**Christine Dyous**
**Michelle Klausner**
**Nancy McGarry**
**Rachel Luck**
**William Rhodes**
Abt Associates

**John Rice, Project Officer**
NCEE/IES

ies NATIONAL CENTER FOR
EDUCATION EVALUATION
AND REGIONAL ASSISTANCE
Institute of Education Sciences

**U.S. Department of Education**
Arne Duncan
*Secretary*

**Institute of Education Sciences**
Sue Betka
*Acting Director*

**National Center for Education Evaluation and Regional Assistance**
Phoebe Cottingham
*Commissioner*

**March 2009**

IES evaluation reports present objective information on the conditions of implementation and impacts of the programs being evaluated. IES evaluation reports do not include conclusions or recommendations or views with regard to actions policymakers or practitioners should take in light of the findings in the reports.

To order copies of this report,

- Write to ED Pubs, Education Publications Center, U.S. Department of Education, P.O. Box 1398, Jessup, MD 20794-1398.
- Call in your request toll free to 1-877-4ED-Pubs. If 877 service is not yet available in your area, call 800-872-5327 (800-USA-LEARN). Those who use a telecommunications device for the deaf (TDD) or a teletypewriter (TTY) should call 800-437-0833.
- Fax your request to 301-470-1244.
- Order online at www.edpubs.org.

This report also is available on the IES website at http://ncee.ed.gov.

**Alternate Formats**
Upon request, this report is available in alternate formats such as Braille, large print, audiotape, or computer diskette. For more information, please contact the Department's Alternate Format Center at 202-260-9895 or 202-205-8113. 8113.

# Contents

# List of Exhibits

# Acknowledgments

# Disclosure of Potential Conflicts of Interest[1]

The research team for this evaluation consists of a prime contractor, Abt Associates Inc., and three subcontractors, Branch Associates, Moore and Associates, and the Center for Resource Management. None of these organizations or their key staff has financial interests that could be affected by findings from the *Impact Evaluation of the U.S. Department of Education's Student Mentoring Program*. No one on the Technical Advisory Panel, convened to provide advice and guidance, has financial interests that could be affected by findings from the evaluation.

---

[1] Contractors carrying out research projects for IES frequently need to obtain expert advice and technical assistance from individuals and entities whose other professional work may not be entirely independent of or separable from the particular tasks they are carrying out for the IES contractor. Contractors endeavor not to put such individuals or entities in positions in which they could bias the analysis and reporting of results, and their potential conflicts of interest are disclosed.

# Executive Summary

This report summarizes the findings from a national evaluation of mentoring programs funded under the U.S. Department of Education's (ED) Student Mentoring Program. The Office of Management and Budget (OMB) requested that the Institute of Education Sciences (IES) within ED oversee an independent evaluation of the Student Mentoring Program. In 2005, ED contracted with Abt Associates and its team of subcontractors, Branch Associates, Moore and Associates, and the Center for Resource Management, to conduct the Impact Evaluation of Student Mentoring Programs. The impact evaluation used an experimental design in which students were randomly assigned to a treatment or control group. Thirty-two purposively selected School Mentoring Programs and 2,573 students took part in the evaluation, which estimated the impact of the programs over one school year on a range of student outcomes. The evaluation also describes the characteristics of the program and the mentors, and provides information about program delivery.

## School-Based Mentoring

School-based or student mentoring programs grow out of interest in youth mentoring generally. The rationale for mentoring is that supportive adults can serve as mentors and can help students avoid high-risk activities and make more successful transitions to adulthood (Sipe, 1996; Tierney and Grossman, 2000; Rhodes, 2002). Youth mentoring programs have emerged as a means to further these goals by connecting at-risk youth with volunteer mentors from outside the family who serve as role models, provide support and guidance, expose students to new things, and provide academic assistance. School- (as opposed to community-) based mentoring programs are programs where typically teachers and other school staff target and identify academically and/or social/emotionally at-risk students whom they feel would benefit from mentoring. These programs then pair these at-risk students with volunteers who meet with them regularly at school (typically one hour per week) either during or after the school day (Portwood and Ayers, 2005).[2] Theoretically, school-based programs also allow mentors and students to focus on academic-related activities such as homework help, tutoring, and reading (Portwood and Ayers, 2005). However, based on prior research findings, programs have been shown to vary widely with regards to the amount of time spent on academics versus social activities (Herrera, Sipe, and McClanahan, 2000; Herrera, Grossman, Kauh, Feldman, and McMaken, 2007).

Over the past several years, school-based mentoring programs have become an increasingly popular way to provide students with mentors (Herrera et al., 2007). This may be due to, at least in part, a number of perceived advantages over community-based mentoring. For example, school-based programs tend to cost less to run per relationship than community-based mentoring programs due to more in-kind contributions from the schools and less overhead (Rhodes, 2002; Portwood and Ayers, 2005; however, see Herrera et al, 2007). However, there are also limitations to the school-based approach. The biggest difference is that school-based mentoring tends to be less intensive than community-based mentoring. For example, the school calendar generally constrains the maximum length of a match to approximately 9 months, which is less than the minimum 12 months of mentoring recommended by those in the mentoring field (e.g., Rhodes, 2002). In practice, the actual

---

[2]    While school-based mentoring is typified by mentors and students meeting on school grounds, it does not exclude mentors and students also getting together at other locations.

length of the school-based mentoring relationship may be even shorter. For example, studies have found a 2- to 3-month time lag from the beginning of the school year in getting students matched with mentors (Hansen, 2005, Herrera, et al., 2000; Karcher, 2008) so that actual mentoring takes place for a period of 5 to 6 months for approximately 6 hours a month (Herrera, et al., 2000). In addition, the school-based approach has often relied heavily on high-school and college-age mentors, which, on one hand, increases the number of students a program can serve, but also can limit the length of the mentoring relationship given the mentors' inability to commit beyond a semester or school year (Herrera et al., 2007). Furthermore, the meta-analytic review of DuBois, Holloway, Valentine, and Harris (2002) regarding the effectiveness of mentoring programs for youth suggests that school-based mentoring programs may be less effective than community-based efforts. In short, compared to community-based mentoring programs, the constraints placed upon school-based mentoring often result in more limited opportunities for students to develop enduring, trusting relationships with adult role models. In turn, school-based mentoring may not be able to provide a sufficient "dosage" of mentoring to achieve lasting positive effects on students.

Research findings on the impacts of school-based mentoring on student outcomes have been limited by weak research designs, small sample sizes, and non-objective measures. However, there is a growing body of more rigorous research that has produced a range of impact findings, generally not sustainable over time. For example, the recent experimental impact evaluation of Big Brothers Big Sisters (BBBS) school-based efforts suggests that school-based programs have the potential to improve students' academic performance, behavior in school, and school attendance (Herrera et al., 2007). These results, however, with the exception of skipping school, did not endure into the following school year. In contrast, a recent experimental evaluation of another school-based mentoring program (the *Study of Mentoring in the Learning Environment* (SMILE)) revealed small, positive effects of mentoring on students' connectedness to peers and on self-esteem and social skills, but not on academic outcomes (Karcher, 2008). Finally, two experimental studies of the Across Ages mentoring program, which has characteristics of both school- and community-based programming (Taylor, LoSciuto, Foz, and Sonkowsko, 1999; Aseltine, Dupre, and Lamlein, 2000), found that the program led to lower levels of student substance use and problem behaviors and stronger attachment of students to school and their families, which were not sustained beyond the end of the school year.

## The Student Mentoring Program

The U.S. Department of Education's Student Mentoring Program, authorized under the No Child Left Behind Act (NCLB) of 2002, Section 4130, is a competitive federal grant program managed by the Office of Safe and Drug Free Schools (OSDFS). It addresses the lack of supportive adults at critical junctures in the lives of students at risk by providing funds to schools and to community- and faith-based organizations to create school-based mentoring programs targeting children in grades 4–8.

The legislation authorizing the program permits program grantees to be responsible for a number of activities including identifying students for the program; recruiting, training and screening of potential mentors (including reference checks and criminal background checks) and supporting mentors through technical assistance and suggested programming. While specific mentoring activities are not mandated in the legislation, the program purpose description states that supported activities are those designed to: improve interpersonal relationships with peers, teachers, other adults and family members; increase personal responsibility and community involvement; discourage drug

and alcohol use, use of weapons and other delinquency involvement; reduce dropout rates; and improve academic achievement.

An absolute priority of the program, as stipulated by OSDFS in their grant solicitation for the program, is its focus on the academic and social needs of at-risk students. In addition to setting the absolute priority, OSFDS, in their grant solicitation, also outlined a number of strategies underlying well-designed and effective school-based mentoring programs including: screening of all potential mentors including background checks; training and support for mentors and program staff on an ongoing basis; activities for mentors and students; and established procedures for supervising and monitoring of mentoring relationships.

## Evaluation Design

This study employs a student-level random assignment design. Specifically, the current evaluation focuses on the impacts of the Student Mentoring Program on students randomly assigned to participate in the ED-funded programs compared to similar students who signed up to participate but were not assigned to participate in the programs.[3] Thus, the study provides experimentally-based evidence about the efficacy of school-based mentoring programs when implemented by a variety of sponsoring organizations.

The key research questions that the evaluation addresses are:

- What is the impact of ED school-based mentoring programs on students' interpersonal relationships with adults, personal responsibility, and community involvement?

- What is the impact of ED school-based mentoring programs on students' school engagement (e.g., attendance, positive attitude towards school) and academic achievement?

- What is the impact of ED school-based mentoring programs on students' high-risk or delinquent behavior?

The sampling pool for this evaluation was based on 255 mentoring programs funded by ED in either 2004 or 2005. The study collected and aggregated data from two cohorts of students: one from the 2005–2006 school year and another from the 2006–2007 school year. The original evaluation design was based on only one cohort. The sample size calculations for this design were based on the assumption that mentoring would be provided to students for an entire school year. When it became apparent after the first program year, however, that the average amount of mentoring was much shorter (i.e., between five and six months) it was decided that in order to conduct a fair test of the program, a larger sample would be needed to detect a smaller effect size. Thus, two cohorts of students were recruited to reach the necessary sample size.

To be selected for the Impact Study, each grantee had to meet three criteria:

---

3    The study is limited to treatment effectiveness across the 32 purposively selected programs, and thus does not generalize outside these programs.

- Be operational so that it could recruit and match students to mentors in the Fall 2005 for the first group of grantees and Fall 2006 for the second group;

- Able to over-subscribe or identify excess demand supporting experimental study needs for an un-served control group (i.e., able to provide tangible evidence of a pool of 4th through 8th grade students referred to the mentoring program) of adequate size to support study requirements; and

- Willing and able to cooperate with the data collection and logistical needs of the national evaluation, including random assignment.

## The Study Sample

A total of 32 unique grantees met the above selection criteria and agreed to participate, comprising the final purposive sample. When identifying students for the study, grantees had categorical criteria to determine eligibility, such as grade level or school location. Sites also identified appropriate students in a variety of ways, most often asking school staff (such as teachers or counselors from the participating schools) to identify and refer students in need of mentoring to the program. To obtain an adequate sample size of students from the 32 grantees, a total of 2,573 students were recruited, 1,272 of whom were randomly assigned to receive mentoring services from the program and 1,301 that were randomly assigned to not receive these services.[4] Students assigned to the control group were free to seek out other mentoring services in the community.

The majority of grantees participating in the Impact Study were non-profit/community-based organizations or faith-based organizations (66 percent) with an average of 6 years of experience with school-based mentoring programs. The average grantee in the Impact Study served 217 students with an annual budget of approximately $277,000. The majority of the grantees in the Impact Study reported having their school-based mentoring programs being extremely focused on improving student academic outcomes (91 percent), increasing students' self-esteem (84 percent), providing students with general guidance (72 percent), and improving students' relationships (63 percent). The majority of students served by the Impact Study grantees were female (57 percent) and a plurality were black or African American (41 percent), and in grades 6 through 8 (44 percent).

In addition to data for the 32 Impact Study grantees, data capturing program characteristics were also collected for a random sample of 100 grantees.[5] The purpose of this random sample was to assess if the purposive sample used to assess program impacts was representative of the full universe of grantees funded through the Student Mentoring Program in 2004 and 2005 for some observable characteristics, as well as to provide additional descriptive information to ED.

Compared to this representative sample of randomly selected grantees, the grantees participating in the Impact Study were less likely to be non-profit/community-based organizations or faith-based organizations but more likely to be school districts, had more years of experience running school-based mentoring, had a larger annual budget, and served more students. Regarding program focus,

---

[4] Of these 32 grantees, 10 provided students in both rounds of recruitment and random assignment. In other words, 10 of the 21 grantees recruited in 2006 had previously been recruited (and were part of our sample) in 2005, leaving us with 32 unique grantees.

[5] A total of 12 out of the 100 randomly sampled grantees were also part of the purposive sample.

grantees in the Impact Study reported being less focused on improving students' academic outcomes and on teaching risk avoidance than the grantees in the representative sample. In addition, there were differences in the students served with grantees in the Impact Study serving more females and more Asian, Latino, and Pacific Islander students but fewer white students than the grantees in the representative sample. The Impact Study sample may also have differed from the representative sample of grantees in other ways that were not observed.

The student sample for the Impact Study had the following characteristics:

- **Gender:** The student sample was 47 percent male versus 53 percent female.

- **Age:** The average age of the sample was 11.2 years old.

- **Race/Ethnicity**: Forty-one percent of the student sample was black or African American, and 31 percent was Hispanic.

- **Poverty**: Eighty-six percent of the sample was eligible for either free or reduced price lunch.

- **Family structure**: Fifty-six percent of the student sample came from two-parent households.

- **Risk status**: Three-fifths (60 percent) of the student sample was at academic risk, defined by being below proficiency in either reading/English language arts (ELA) or math (or both) at baseline, and one-fourth of the sample (25 percent) was at risk for delinquency, defined by self-reported delinquent behaviors.

- **Prior mentoring experience**: Twenty-six percent of the sample reported receiving mentoring in the prior school year.

- **Of the baseline characteristics assessed, only one statistically significant difference between the treatment and control group was observed.** A higher proportion of students in the treatment group were eligible to receive free or reduced-price school lunches than in the control group.

## Program Delivery Findings

Both grantees and mentors were surveyed to describe various characteristics of program delivery, including training and support for mentors, characteristics of mentors, matching of students with mentors, and mentor/student relationship duration and activities. On average, grantees in the study implemented the program following the guidance provided by the legislation and program office. Also, program delivery was, by and large, consistent with findings from previous studies of school-based mentoring.

Key findings on program delivery include the following:

- **Approximately one in ten mentors reported not having undergone a reference or background check despite being required by the program as a condition of the grant.** Eleven percent of mentors reported not having had either a background or reference check conducted pre-match, despite the fact that all 32 grantees indicated requiring some form of background screening before matching mentors with students. Because only mentors and not grantees were asked this question, it is possible that some mentors were simply unaware (or had forgotten) that a background or reference check was conducted by the grantee.

- **The majority of mentors received pre-match training or orientation and had access to ongoing supports from the program.** Ninety-six percent of mentors reported receiving an average of 3.4 hours of some form of pre-match training or orientation. Forty-one percent of mentors reported that ongoing training was available after they had begun meeting regularly with their students. Ninety-four percent reported having access to some kind of ongoing supports, consistent with legislative and program guidelines.

- **The majority of students were matched with mentors of the same race and gender.** Fifty-five percent of matches in our study were between individuals who had the same racial status. Eighty-one percent of matches in our study were between students and mentors of the same gender.

- **The majority of mentors met with their students on a one-to-one basis.** Mentors, on average, also reported meeting with their students, on a weekly basis for approximately one hour per meeting. This finding is consistent with findings from other studies (Herrera et al. 2007; and Karcher, 2008).

- **Seventeen percent of the students randomly assigned to the treatment group never received mentoring from the program.** This includes 14 percent of students in the treatment group who were never matched with mentors and another 3 percent who were matched with mentors, but never actually met. However, the percentage of unmatched students in this study is within the range of past experience engaging mentors in randomized impact studies of mentoring.[6]

- **On average, the programs took a total of 81 days to match students and mentors, from the start of the school year.** On average, there was a lag of 37 days between the date of random assignment and the time when the student was matched. This lag between the beginning of the school year and matching students is consistent with findings from previous research (Herrera et al, 2000, Hansen, 2005, and Karcher, 2008).

- **For students who were matched and met with their mentors, the average length of the relationship was 5.8 months.** This finding, however, is consistent with previous research. For example, Herrera et al. (2007) in the impact study of the BBBS school-based mentoring program reported an average match length of 5.3 months.

---

[6] One school-based mentoring study and one community-based mentoring study (both random assignment of students to conditions) has found that the proportion of students slated to receive mentoring services that remain unmatched with mentors was 7 and 22 percent, respectively (Herrera, et al., 2007; Tierney and Grossman, 2000).

- **Discussing relationships and future plans, and to a lesser extent, working on academics were the most frequent activities reported by mentors.** Mentors and students worked together on a range of activities. Approximately half of the mentors reported frequently discussing relationships and future plans (52 percent and 48 percent, respectively). In contrast, 43 percent reported working frequently on academics, while 21 percent reported <u>never</u> working on academics. The greater focus on the social needs of the students compared to academic needs has been found in some, but not all, of the previous research.[7]

- **Approximately 20 percent of the mentors were of high-school age (18 years or younger) and an additional 23 percent were of college-age.** However, this is still a smaller percentage than findings from previous research where the majority of mentors were of high school or college age.[8]

## The Treatment Contrast

- **Eighty-six percent of treatment group students reported receiving mentoring services (through any program) over the past school year compared to 35 percent of the control group students.** Treatment group students received services from ED Student Mentoring programs and control group students reported receiving mentoring either from the program or elsewhere in the community.[9]

- **Students in the treatment group reported receiving more intensive mentoring than students in the control group who received mentoring.** Eighty-five percent of students in the treatment group who actually received mentoring reported meeting with their mentors at least twice a month, compared to 66 percent of the mentored control group students.

## Outcome Measures

We measured a total of 17 outcomes in the domains of interpersonal relationships and personal responsibility, academic achievement and engagement, and high-risk or delinquent behavior. We based the evaluation outcome measures on the intended outcomes as stated in the legislation authorizing the program and the three research questions. Using both self-report data from the students and school records, the study estimated impacts on outcomes for treatment and control students, based on data from two time points—in the fall of the school year and at the end of the school year.

---

[7] Less than a third of the mentors in the BBBS study (Herrera et al., 2007) reported spending a lot or most of their time on academic activities.

[8] For example, Herrera et al. (2007) reported that half of the mentors in the BBBS study were 18 years old or younger, with an additional 17 percent 19 to 24 years old. In Karcher's 2008 study of school-based mentoring, 70 percent of the mentors were college students.

[9] Three percent of students in the control group received mentoring from the grantees in the study.

## Analytic Approach

The analysis strategy utilized a fixed-effects model to estimate the average treatment effect across all programs for students assigned to receive mentoring versus students assigned to an untreated control group.[10] The fixed effects model was also used to examine five subgroup differences: (1) gender, (2) age (students 12 or older versus students less than 12 years old), (3) family structure (students from two-parent families versus students from other types of families), (4) presence of self-reported delinquent behaviors at baseline (theft, possession of a weapon, drug use, alcohol use, or gang activity), and (5) academic non-proficiency (in math, reading/English Language Arts (ELA), or both) at baseline. We obtained impact estimates for each of the selected subgroups using the same approach as in the main analysis. We then performed a t-test to identify any statistically significant differences in impacts between each paired set of subgroups – for example, to test whether the estimated impact of school-based student mentoring on boys was different from the impact on girls in our sample. To control for chance findings, a multiple comparisons procedure, known as the Benjamini-Hochberg (BH) correction, was employed within each outcome domain in analysis of the full sample and within each outcome domain in each of the five subgroup analyses.

Finally, given that characteristics of programs and their mentors varied considerably across sites, we wished to determine whether some sites or groups of sites could be characterized as more or less successful, and, if so, whether we could identify program characteristics associated with differences in impacts at the site level. Therefore, a series of exploratory analyses were also conducted to explore site-level differences in impacts.

## Impacts of the Student Mentoring Program

### Estimation of Overall Impacts of the Student Mentoring Program

We estimated a total of 17 impacts in three domains: (1) academic achievement and engagement; (2) interpersonal relationships and personal responsibility; and (3) high-risk or delinquent behavior.

- **The Student Mentoring Program did not lead to statistically significant impacts on students in any of the three outcome domains.** The estimated impact on the Student Mentoring Program on the outcome measures for all three domains is reported in Exhibit ES.1.

- Three of the impacts were statistically significant before accounting for multiple comparisons. However, after accounting for multiple comparisons within each of the three domains, these three impact estimates were no longer statistically significant.

---

[10] We use the term "fixed-effects" within the dual perspectives of sampling and statistical inference. Because student mentoring programs were chosen purposively, not randomly into the study, results cannot be generalized to the full universe of programs. The fixed-effects model is therefore, appropriate, given our level of inference does not extend beyond our study sample of purposively chosen programs.

**Exhibit ES.1: Estimated Overall Impacts on Student Outcomes**

| Outcome | Unadjusted Mean Outcome | | | | Estimated Impact | | | |
|---|---|---|---|---|---|---|---|---|
| | Treatment Group | | Control Group | | Regression Adjusted T-C Group Difference [a] | p-value to Test Difference | BH-Corrected Critical Value [b] | Estimated Effect Size |
| | Mean | Standard Deviation | Mean | Standard Deviation | | | | |
| **Interpersonal Relationships, Personal Responsibility, and Community Involvement** | | | | | | | | |
| **Self-Reported Outcome (Range 1- 4)** | | | | | | | | |
| Pro-Social Behaviors | 2.79 | 0.81 | 2.80 | 0.80 | -0.01 | 0.68 | NA | -0.01 |
|     Percent Missing Data | ≤3% | | ≤3% | | | | | |
| **Academic Outcomes** | | | | | | | | |
| **Self-Reported Outcome (Range 1- 4)** | | | | | | | | |
|   School Efficacy and Bonding | 3.06 | 0.80 | 3.03 | 0.85 | 0.04 | 0.08 | 0.02 | 0.07 |
|   Future Orientation | 3.85 | 0.54 | 3.80 | 0.63 | 0.03* | 0.04 | 0.01 | 0.08 |
|     Percent Missing Data | ≤3% | | ≤3% | | | | | |
| **School-Reported Outcome** | | | | | | | | |
| ***Overall Absenteeism Rate (all absences as percent of total days enrolled)*** | 5.03 | 7.71 | 5.49 | 9.63 | -0.46* | 0.04 | 0.01 | -0.09 |
|     Percent Missing Data | 15% | | 18% | | | | | |
| ***Grades (Range 1–5)*** | | | | | | | | |
| Math | 3.19 | 1.70 | 3.23 | 1.67 | -0.05 | 0.23 | 0.02 | -0.05 |
| Reading/ELA [c] | 3.57 | 1.78 | 3.61 | 1.69 | -0.04 | 0.40 | 0.03 | -0.04 |
| Science | 3.52 | 1.87 | 3.55 | 1.86 | -0.03 | 0.48 | 0.04 | -0.03 |
| Social Studies | 3.53 | 1.92 | 3.56 | 1.83 | -0.01 | 0.78 | 0.05 | -0.01 |
|     Percent Missing Data | ≤35% | | ≤33% | | | | | |
| ***State Assessment Tests*** | | | | | | | | |
| Math—Percent Proficient | 45.69 | | 47.10 | | -1.53 | 0.41 | 0.04 | 0.94[d] |
| Reading/ELA—Percent Proficient | 49.40 | | 50.76 | | -1.67 | 0.37 | 0.03 | 0.94[d] |
|     Percent Missing Data | ≤25% | | ≤20% | | | | | |
| **Delinquent Behaviors and Participation in Harmful Activities** | | | | | | | | |
| **Self-Reported Outcome (Range 1–4)** | | | | | | | | |
|   Misconduct [e] | 3.20 | 0.86 | 3.20 | 0.85 | 0.00 | 0.95 | 0.05 | 0.00 |
|   Delinquency [e] | 3.87 | 0.36 | 3.85 | 0.40 | 0.01 | 0.28 | 0.02 | 0.04 |
|     Percent Missing Data | ≤3% | | ≤4% | | | | | |
| **School-Reported Outcome** | | | | | | | | |
| ***Truancy (unexcused absences as percent of total days enrolled [f])*** | 2.04 | 4.8 | 2.47 | 6.91 | -0.45* | 0.02 | 0.01 | -0.14 |
| ***Misconduct*** | | | | | | | | |
| Percent committing any infraction | 25.00 | | 22.91 | | 2.56 | 0.13 | 0.01 | 1.59[d] |
| Percent committing repeated infractions (2+) | 14.21 | | 15.63 | | -0.98 | 0.48 | 0.04 | 0.93[d] |
| ***Delinquency*** | | | | | | | | |
| Percent committing any infraction | 18.13 | | 20.03 | | -1.51 | 0.35 | 0.03 | 0.91[d] |
| Percent committing repeated infractions (2+) | 8.64 | | 9.13 | | -0.56 | 0.65 | 0.04 | 0.93[d] |
|     Percent Missing Data | ≤22% | | ≤23% | | | | | |
| Number of students | 1163 | | 1197 | | | | | |

[a]    Regression Adjusted T-C Group Difference will not necessarily be equal to the difference between the Unadjusted Mean Outcomes.

[b]    Based on Benjamini-Hochberg test.

[c]    ELA = English/Language Arts.

[d]    Odds-ratio.

[e]    Higher scores on the Self-Reported Misconduct and Delinquency scales indicate more positive outcomes.

[f]    Based on 27 sites that reported unexcused absences and total days enrolled. Treatment Group: Missing data ≤38%; Control Group: Missing data ≤36%.

\*    p-value (of estimated impact) < 0.05, two-tailed test.

+    p-value (of estimated impact < BH-Corrected Critical Value ➔ statistically significant holding the false discovery rate under multiple testing below 0.05

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Estimation of Subgroup Effects**

Several subgroup analyses were statistically significant after accounting for multiple comparisons.

- **The Student Mentoring Program improved academic outcomes for girls and produced mixed academic outcomes for boys.** There were several positive impacts of the program for girls. The impact on self-reported scholastic efficacy and school bonding was positive and statistically significant for girls, with treatment group girls scoring higher than control group girls. In addition, there was a statistically significant difference in impacts on the scholastic efficacy and school bonding measure by gender (effect size for girls = 0.18 versus -0.05 for boys). There was also a positive, statistically significant effect on future orientation for boys (effect size = 0.17). However, the difference in impacts between boys and girls on this measure was not statistically significant.

- **For boys, the Student Mentoring Program negatively affected self-reported pro-social behavior** Boys who were assigned to mentoring reported statistically significant lower scores on the pro-social behaviors scale compared to their control group peers. Moreover, there was a statistically significant difference in impacts on the pro-social behaviors scale by gender (effect size for girls = 0.08 versus – 0.11 for boys). .

- **The Student Mentoring Program led to a decrease in truancy for younger students.** Truancy (i.e., unexcused absence) showed a statistically significant improvement for younger students (below age 12) who were assigned to mentoring compared to same age peers in the control group (effect size = -0.23). However, the difference in impacts on truancy between younger and older students (aged 12 and older) was not statistically significant after accounting for multiple comparisons.

**Site-Level Characteristics and Impacts**

Although we did not find that the Student Mentoring Programs had statistically significant impacts on student-level outcomes for our sample as a whole, we wished to determine whether characteristics of programs and their mentors varied across sites and, if so, whether we could identify program and mentor characteristics associated with differences in impacts at the site level. Because sites were not randomly assigned to different levels of implementation—a primary potential source of impact variation—this analysis is descriptive and exploratory, not causal, in nature.

For this analysis, it was essential to develop a parsimonious model for testing for any relationship between program and mentor characteristics (and contextual factors) and site-level impacts. Therefore, in choosing the final set of site-level covariates for inclusion in our model, we considered several factors, including their theoretical importance in influencing impacts, possessing statistically significant site-level variation, and low site-level correlations among these variables to avoid problems with multicollinearity.[11]

The site-level covariates in our analysis included nine factors: (1) average hours of pre-match training provided to mentors; (2) amount of ongoing mentor support (average frequency of mentor-supervisor

---

[11]    In general, we focused on proximal factors thought to have a direct influence on impacts rather than distal factors, which may be correlated with impacts, but whose influence may be indirect and/or mediated through more proximal causes.

meetings); (3) use of activities in mentor/student meetings (e.g., percent of mentors reporting almost always/most of the time either working on homework and/or academic skills with students); (4) percent of mentors aged 22 or below; (5) percent of mentor/student matches of the same race/ethnicity; (6) percent of students with self-reported delinquent behaviors at baseline; (7) percent of students scoring "not proficient" in either math or reading/ELA at baseline; (8) percent of mentor/student matches lasting 6 months or longer; and (9) average total hours of mentor/student meetings per month.[12]

Although we did not explicitly control for multiple comparisons because this was an exploratory analysis, it is important to note that we conducted 153 individual hypothesis tests of potential associations between the 9 covariates and the 17 outcome measures, for roughly 7 or 8 of which we would expect to reject the null hypothesis at the 0.05 level by random chance alone. In fact, we found 12 statistically significant relationships.

The following associations between site-level impacts and each of these site characteristics were statistically significant at the 95 percent confidence level, holding all other characteristics constant:[13]

- **The frequency of mentor/supervisor meetings was negatively associated with site-level impacts.** All other things equal, the frequency of mentor/supervisor meetings was negatively associated with site-level impacts on the Pro-social Behaviors measure from the Student Survey and on grades in math and social studies. They were also positively associated with site-level impacts on school-reported delinquency.

- **The proportion of students with self-reported delinquent behaviors at baseline had both positive and negative relationships with site-level impacts**. The proportion of students with self-reported delinquent behaviors at baseline was positively associated with site-level impacts on social studies grades and negatively associated with site-level impacts on absenteeism and truancy.

  However, the proportion of students with self-reported delinquent behaviors at baseline was also positively associated with site-level impacts on repeated misconduct from student records.

- **The proportion of mentors aged 22 or younger was negatively associated with site-level impacts on math grades.**

- **The proportion of mentor/student matches of the same race/ethnicity was positively associated with site-level impacts on reading/ELA grades.**

- **Average monthly hours of mentor/student meetings had both positive and negative relationships with site-level impacts.** Average monthly hours of meeting were

---

12  We also included in our analyses an indicator variable for the share of the control group that received mentoring (from any source) during the outcome period to adjust for potential differential attenuation of impact estimates from site to site.

13  For the purposes of reporting associations between site-level characteristics and impacts, we refer to relationships as "positive" or "negative" in the statistical sense, reflecting the direction of the coefficient. However, in some cases a positive statistical relationship denotes a negative substantive relationship or a negative statistical relationship denotes a positive substantive relationship.

positively associated with site-level impacts on student self-reported future orientation, but negatively associated with site-level impacts on grades in math and reading/ELA.

# Chapter 1: Overview and Background of Study

## 1.1 Introduction

This report summarizes the findings from a national evaluation of mentoring programs funded under the U.S. Department of Education's (ED) Student Mentoring Program. The Office of Management and Budget (OMB) requested that the Institute of Education Sciences (IES) within ED oversee an independent evaluation of the Student Mentoring Program. In 2005, IES contracted with Abt Associates and its team of subcontractors, Branch Associates, Moore and Associates, and the Center for Resource Management (CRM), to conduct the Impact Evaluation of Student Mentoring Programs. This three-and-one-half-year evaluation is designed to describe ED student mentoring programs and to estimate the short-term impact (i.e., over the period of one school year) of these programs on a range of student outcomes.

The impact evaluation used an experimental design in which students were randomly assigned to a treatment or control group. Two groups of ED funded mentoring programs defined the sampling pool for this evaluation—165 grantees funded in 2004 and an additional 90 grantees funded in 2005.[14] Data for the impact evaluation were drawn from the 32 grantees willing and able to randomly assign students who were interested and eligible to receive mentoring to either receive or not receive program services.[15] In addition, data capturing program characteristics were collected for each of these 32 purposively selected programs, as well as for a random sample of 100 grantees.[16] The purpose of this random sample was to determine if the purposive sample used to assess program impacts was comparable to the population of ED's program grantees for some observable characteristics, as well as to provide additional descriptive information to ED for program improvement purposes.

This chapter presents a general overview of the study including a description of school-based mentoring programs and in particular, ED's Student Mentoring Program, a short review of prior research on school-based mentoring programs, and a brief overview of the evaluation, including the key research questions. The remainder of this report provides further information on the study sample and design, and discusses the evaluation findings, both in terms of program implementation and impacts on students. Chapter 2 discusses the study sample and design, covering topics such as statistical power and construction of outcome measures. Chapter 3 provides a discussion of the program implementation findings, from the perspective of both the grantees and mentors. Chapter 4 discusses the evaluation findings from the perspective of impacts on students, including overall and for subgroups. A discussion of the relationships between program characteristics and site-level impacts is also presented in this chapter. Further technical material is contained in Appendices A–G.

---

[14] Although 2004 and 2005 were the first years of ED funding for all of these 255 programs, some of these programs had already been in operation prior to receiving ED funds. This fact, however, had no bearing on each program's potential participation in this evaluation because students already receiving services from newly funded programs already in operation were not eligible to be included in the study.

[15] Of these 32 grantees, 10 provided students in both rounds of recruitment and random assignment. That is, 10 of the 21 grantees recruited in 2006 had previously been recruited (and were part of our sample) in 2005, and an additional 11 grantees were recruited in 2006, leaving us with 32 unique grantees.

[16] A total of 12 out of the 100 randomly sampled grantees were also part of the purposive sample.

## 1.2  School-Based Mentoring

The Student Mentoring Program is a school-based effort.  It grows out of interest in student mentoring generally.  The rationale for mentoring is supported by research from a variety of fields that suggests that supportive adults serving as mentors can help students avoid high-risk activities and make more successful transitions to adulthood (Sipe, 1996; Tierney and Grossman, 2000; Rhodes, 2002).[17] This type of intervention may be particularly helpful for students from single-parent families and families in poverty (Lee and Cramond, 1999). Factors associated with those types of families have made it increasingly difficult for economically disadvantaged students to connect with "natural mentors," or adults from their neighborhoods or social networks with whom mentoring relationships could evolve organically (DuBois and Silverthorn, 2005; Eccles and Gootman, 2002).  Mentoring programs have emerged in response to this problem to connect at-risk students with volunteer mentors from outside the family who serve as role models, provide support and guidance, expose students to new things, and provide academic assistance.

School- (as opposed to community-) based mentoring programs are programs in which targeted students are paired with volunteers who meet with them regularly *at school* either during or after the school day.[18]  Over the past several years, school-based mentoring programs have become an increasingly popular way to provide students with mentors (Herrera et al., 2007).  School-based mentoring programs are distinguished from other mentoring programs not only on the basis of where the mentoring activities occur (i.e., in the school setting), but also in how the student participants are identified (by school personnel). Teachers and other school staff identify students whom they feel would benefit from mentoring, and students and mentors often meet during school hours or as part of an after-school program, rather than in the community.[19] In addition, school-based mentoring may also differ from community-based programs in terms of activities conducted during the mentoring sessions. It should be noted that the meta-analytic review of DuBois, et al. (2002) of the effectiveness of mentoring programs for youth suggests that school-based mentoring programs may be less effective than community-based efforts.

Most school-based mentoring programs foster one-to-one relationships and encourage mentors and students to commit to working together from their initial meeting through at least the end of the school year. Although they take place at schools, school-based mentoring programs are not necessarily expected to focus exclusively on improving students' academic performance or school attachment; they can also focus on other outcomes as well: improving interpersonal relationships,

---

[17]  Although most of these studies were not explicitly about school-based mentoring, the theory supporting school-based mentoring rests on many of the same assumptions as mentoring in general.

[18]  In community-based mentoring programs, students and mentors meet outside of school grounds. Community-based programs also tend to require more pre-screening activities (because they are less likely to be supervised), may foster longer-lasting relationships (because the relationships happen outside of the school year) and may be less likely to use college or high school students as volunteers. Community-based mentoring programs also tend to cost more to run per relationship than do school-based mentoring programs. For more on the differences between community- and school-based mentoring programs, see Portwood and Ayers (2005).

[19]  While school-based mentoring is typified by mentors and students meeting on school grounds, it does not exclude mentors and students also getting together at other locations.

personal responsibility and community involvement; and reducing juvenile delinquency and participation in harmful activities.

## 1.3 Description of the Student Mentoring Program

**Legislative and Programmatic Framework**

The U.S. Department of Education's Student Mentoring Program, authorized under the No Child Left Behind Act (NCLB) of 2002, Section 4130, is a competitive federal grant program managed by the Office of Safe and Drug Free Schools (OSDFS). It addresses the lack of supportive adults at critical junctures in the lives of students at risk by providing funds to schools, nonprofits/community-based organizations, and faith-based organizations to create school-based mentoring programs targeted at grades 4–8. **An absolute priority of the program, as stipulated by OSDFS in their grant solicitation for the program, is its focus on the academic and social needs of at-risk students.**

As of this date, ED has funded several hundred grantees with approximately $204 million to implement the program. As mandated in the NCLB legislation, grantees are responsible for identifying students for the program; recruiting, screening and training mentors; matching mentors and students; and supporting and monitoring relationships to ensure that they benefit targeted students. Mentors assist students by providing mentoring to students in grades 4-8 at risk for poor academic outcomes, dropping out of school, delinquency and/or gang involvement. They provide general guidance, serve as role models and/or provide academic assistance and encouragement.

**Theory of Action for the Student Mentoring Program**

The theory behind ED's Student Mentoring Program is that facilitating mentoring relationships will result in important long-term impacts for students served, specifically: improved academic outcomes and participation; improved interpersonal relationships, personal responsibility and community involvement; and reduced juvenile delinquency and participation in harmful activities. Exhibit 1.1 presents a logic model of the intervention. The goals of the program and its target population are shown in Box 1, derivative activities to address the goals in Box 2, and impacts that might be expected from those activities in Box 3. In this section, we describe the model in more detail.

**Exhibit 1.1: Logic Model of ED's Student Mentoring Program**

**Program Goals (1)**

Absolute Priority: Address academic and social needs of students with the greatest need.

Develop programs to provide mentoring services for students in grades 4-8 who are at increased risk for poor academic outcomes, dropping out of school, and delinquency and gang involvement. The following groups of students are the program target population:
- Students who lack strong adult role models
- Students living in rural and/or high-crime areas
- Students who have troubled home lives
- Students who have academic challenges
- Students involved in criminal/delinquent activities

**Activities (2)**

**Grantee Activities**

Fund local educational agencies, nonprofits/community-based organizations, and partnerships thereof to, working through schools:
- Identify students at risk
- Recruit, screen, and train mentors
- Match mentors and students in close (1:1 preferred) relationships
- Provide/identify space in the school or other setting for mentoring
- Support and monitor relationships to ensure that they benefit students

**Mentor Activities**

Provide students with support and general guidance

Serve as role models:
- Promote personal and social responsibility
- Expose students to new experiences and examples of opportunity that enhance students' ability to become responsible adults
- Encourage participation in community service and community activities

Provide students with academic assistance and encourage graduation from secondary school and planning for post-secondary education and training

**Impacts on Students (3)**

**Improved Interpersonal Relationships, Personal Responsibility, and Community Involvement**

Improved relationships between students and
- Their peers
- Family members
- Teachers
- Other adults

Increased personal responsibility
Increased civic engagement/volunteering

**Improved Academic Engagement and Achievement**
- Increased participation in elementary and secondary education
- Improved academic achievement
- Improved school attendance
- Reduced dropout rates*
- Improved planning for the future

**Reduced Juvenile Delinquency/ Participation in Harmful Activities**
- Reduced use of drugs/alcohol
- Reduced use of dangerous weapons
- Reduced gang involvement
- Reduced other criminal activities
- Reduced promiscuous behaviors**

---

* Given the targeted age group of this study (grades 4–8), we were not in a position to measure dropout rates.

** While dealing with promiscuity is a goal of the NCLB programming, we did not include an assessment of this outcome in this study. In the initial phases of instrument development we found that questions regarding sexual behaviors or attitudes were not acceptable to principals and parents, particularly for students at the elementary school level.

*Purpose/Goals of the Program (Box 1)*
The Student Mentoring Program provides funding for grantees to develop programs to assist students who lack positive role models and are at risk of educational failure, dropping out, or involvement in criminal or delinquent activities. The program matches at-risk students with trained adult or peer (secondary school) mentors to form positive relationships through regularly scheduled academic, social and recreational activities. ED funding decisions give **absolute** priority to programs that address the academic and behavioral problems of students with "the greatest need" living in rural and/or high crime areas, in troubled home environments, and/or who attend schools with violence problems. The program focuses on students in grades 4 through 8. While the legislation does not limit funding solely to school-based efforts, it gives priority to school-based mentoring programs.

*Activities (Box 2)*
The Student Mentoring Program funds two types of activities: (1) Grantee activities in identifying, training and monitoring mentors; and (2) Mentor activities with students.

The listings in Box 2 under "Grantee Activities" are specified in the legislation. These activities include identifying students for the program; recruiting, training and screening of potential mentors (including reference checks and criminal background checks) and supporting of mentors through technical assistance and suggested programming. Other grantee activities include hiring, training and professional development of mentoring coordinators and support staff, and the development and dissemination of materials and supplies.

The second category of activities represents those specified for participating mentors (Mentor Activities). Mentors in the Student Mentoring Program are adults or high school or college students who volunteer to meet with students on a regular basis throughout the school year. While specific activities are not mandated in the legislation, the program purpose description states that supported activities are those designed to: improve interpersonal relationships with peers, teachers, other adults and family members; reduce dropout rates; discourage drug and alcohol use, use of weapons and other delinquency involvement; and improve academic achievement.[20]

Because of the absolute priority all grantees must focus on both academic and social needs of the mentored students, although grantees have discretion in their specific program objectives and may emphasize some objectives in their choice of activities over others. Also, some programs use standard activities and fixed curricula while others simply provide suggested activities and guidelines for mentors to use. In all cases, the legislation specifies that activities and mentor supports should be designed to ensure longstanding "one to one" relationships (i.e., the mentor is committed for the entire school year and, wherever possible, mentors only one student).

Prior research points to several characteristics of programming that may moderate the effects found in school-based mentoring (DuBois et al., 2002; Grossman, Baldwin, and Johnson, 1999; Herrera, 2004). These include: support and ongoing training of mentors, providing activities, establishing expectations of frequency of contact, support and involvement of parents, length of the match (school year versus longer), and activity focus (academic versus social/emotional).

---

[20]  U.S. Department of Education, Office of Safe and Drug-Free Schools, Mentoring Programs: FY 2004 Information and Application Procedures.

*Intended Impacts on Students (Box 3)*

The expected impacts of the Student Mentoring Program are improvement in the attitudes and behaviors of students in three primary outcome areas: (1) interpersonal relationships, personal responsibility and community involvement; (2) academic performance and participation; and 3) delinquency and other harmful behaviors. The programs funded under this program may work to improve outcomes in one or more specified areas:

- Improved interpersonal relationship with adults, peers and family members;
- Increased personal and social responsibility;
- Increased participation in community services and activities;
- Increased goal setting and planning for the future, including planning for graduation and/or postsecondary education or training;
- Increased participation and improved performance in school; and
- Reduced violence, use of weapons, drugs and alcohol, and other harmful activities.

Research findings on the impacts of school-based mentoring on student outcomes vary considerably across studies. For example, the recent impact evaluation of Big Brothers Big Sisters (BBBS) school-based efforts suggests that school-based programs have the potential to improve students' academic performance, behavior in school, and school attendance (Herrera et al., 2007). Most reported positive results, however, did not endure into the following school year. In contrast, a recent evaluation of another school-based mentoring program (the *Study of Mentoring in the Learning Environment* [SMILE]) revealed small, positive effects of mentoring on students' connectedness to peers and on self-esteem and social skills, but not on academic outcomes (Karcher, 2008). Further analysis of the impact of mentoring on specific subgroups of students indicated additional benefits for both elementary school boys and high school girls, but showed negative effects for older boys and younger girls (Karcher, 2008).

Studies reporting on the effect of school-based mentoring on delinquency and drug use yield similarly mixed results. Two experimental studies of the Across Ages mentoring program (Taylor et al., 1999; Aseltine et al., 2000), which has characteristics of both school- and community-based programming, found that the program led to lower levels of student substance use and problem behaviors and stronger attachment of students to school and their families. However, these benefits were not sustained beyond the end of the school year. In contrast, an evaluation of the YouthFriends school-based mentoring program (Portwood, Ayers, Kinnison, Waris, and Wise, 2005) reported statistically significant but modest differences between mentored students and a matched comparison group on sense of belonging to the school community, but no differences on delinquency, drug and alcohol use, self-esteem, self-concept, goal setting, or attitudes toward adults. In short, variation in focus and measures among previous studies examining school-based mentoring have resulted in a set of inconsistent findings on a range of outcomes in both academic and behavioral domains. Moreover, in general, where positive outcomes have been found, they have either not been followed-up on or have not been sustained past the end of the school year in which mentoring was offered.

## 1.4 Key Research Questions

The design of the evaluation of the Student Mentoring Program, in particular the research questions to be addressed, has been influenced by a growing body of research on the efficacy of mentoring in general, and school-based mentoring in particular (Sipe, 1996; Taylor et al., 1999; Aseltine et al., 2000; Tierney and Grossman, 2000; Rhodes, Grossman and Resch, 2000; Thompson and Kelly-Vance, 2001; Portwood et al., 2005, Herrera et al., 2007; Karcher, 2008). Despite limitations common to many of these studies, such as non-blind teacher evaluation of students, limited comparison groups, small sample sizes, and lack of statistical corrections for multiple outcome measures, the body of accumulated research does provide a basic framework to guide the evaluation of the program.

The evaluation of the Student Mentoring Program builds upon this body of research by providing experimentally-based evidence about the efficacy of school-based mentoring programs when implemented by a variety of sponsoring organization. Specifically, the current evaluation focuses on the impacts of the Student Mentoring Program on students randomly assigned to participate in the ED-funded programs compared to similar students who signed up to participate but were not assigned to participate in the programs.[21] It was designed to address these key research questions:

- What is the impact of ED school-based mentoring programs on students' interpersonal relationships with adults, personal responsibility, and community involvement?

- What is the impact of ED school-based mentoring programs on students' school engagement (e.g., attendance, positive attitude towards school) and academic achievement?

- What is the impact of ED school-based mentoring programs on students' high-risk or delinquent behavior?

In addition to the main focus on the overall impacts of the Student Mentoring Program, the study also examines impacts between subgroups of students along the following dimensions:[22]

- Gender;

- Age;

- Family structure;

- Presence of self-reported delinquent behaviors at baseline; and

- Academic non-proficiency (in math and/or reading/ELA) at baseline.

---

[21] The study is limited to treatment effectiveness across these selected programs, and thus does not generalize outside these programs.

[22] Although the main focus of these subgroup analyses was on the differences in impacts between subgroups of students (e.g., boys versus girls), the study also reports on the separate impacts for each subgroup (e.g., treatment group boys versus control group boys).

Finally, the study presents a set of exploratory analyses examining relationships between site-level characteristics and site-level impacts. The site-level characteristics examined were:

- Average hours of pre-match training provided to mentors
- Percent of mentors aged 22 or below
- Percent of mentor/student matches of the same race/ethnicity

- Amount of ongoing mentor support
- Frequency of use of activities in mentor/student meetings
- Percent of mentor/student matches lasting 6 months or longer
- Average total hours of mentor/student meetings per month.
- Percent of students with self-reported delinquent behaviors at baseline
- Percent of students scoring "not proficient" in either math or reading/ELA at baseline

In addition to providing information about the Student Mentoring Program's impact on targeted student outcomes, the study describes program characteristics and program implementation. It also assesses the extent to which grantees in the study were representative of the full universe of grantees funded through the Student Mentoring Program in 2004 and 2005.

Accordingly, the study was designed to answer the following questions about program characteristics through surveys of grantees and mentors:

- What kinds of organizations implemented the Student Mentoring Program and how much experience did they have running mentoring programs?

- What were the shared characteristics of Student Mentoring Programs?

- What were the characteristics of students served in the Student Mentoring Program?

- What were the characteristics of mentors serving in the Student Mentoring Program?

- What training did mentors receive and what type of ongoing support was provided?

- What was the process for matching students to mentors?

- How often did students meet with their mentors and for what duration?

- In what types of activities did mentors and students participate?

- What kinds of relationships did the Student Mentoring Program cultivate between targeted students and mentors?

- Did students receive mentoring in addition to the mentoring provided by the grantees in the study?

# Chapter 2: Sample Selection, Study Design, and Analytic Approach

## 2.1 Overview of the Design

The evaluation of the Student Mentoring Program rests on an experimental design that involves a two-level sample: a purposive sample of 32 ED-funded programs[23] throughout the country and a sample of 2,573 students in grades 4–8 who were randomly assigned either to a treatment condition (offered mentoring services through these programs) or to a control condition (offered no program mentoring services).

The outcomes of interest include better relationships with parents and other adults, increased personal and social responsibility, increased participation in community services and activities, improved positive attitudes toward school, better school performance, avoidance of risk behaviors and reduced delinquency and misconduct. Using both self report data from the students and school records the study evaluated these outcomes for treatment and control students at two points—in the fall of the school year (i.e., baseline) and at the end of the school year (i.e., follow-up). In addition, mentors for all students in the study were surveyed and data about those mentors are included in the analysis of student level impacts. The analysis strategy estimated the average treatment effect of a number of outcomes across all programs for students randomly assigned to ether a treatment or control condition, including an examination of subgroup differences in each site.

In the following sections we describe in greater detail the evaluation design of the study, covering the following key topics:

- Site recruitment
- Identification and random assignment of student participants
- Statistical power
- Selection of a random sample of comparison grantees
- Grantee and mentor measures
- Measuring student outcomes
- Response rates
- Approach to the analysis of student impacts

## 2.2 Site Recruitment

In this study, a "site" refers to a Student Mentoring grantee comprised by a school district, community-based organization, or faith-based organization providing mentoring to students in one or more schools. At the time of site recruitment, the U.S. Department of Education School-Based Mentoring programs comprised a total of 255 individual grantees: 165 grantees awarded funds in 2004, and 90 grantees awarded funds in 2005. Of these 255 grantees, 10 programs were determined to

---

[23] Students were sampled in two phases or cohorts, in Year 1 and Year 2. The final sample of participating programs numbered 32, but included 10 programs that continued with a new set of students in Year 2 for a total of 42 "program" data points in the analysis.

be not of interest to the study because either (1) they were not operating a student mentoring program; (2) they did not serve the population of students as described in the Student Mentoring Program legislation; or (3) program application data were unavailable.

The remaining pool of 245 grantees constituted the total population eligible for our study, representing the universe of ED-funded grantees operating under Student Mentoring Program guidelines in those years.  To be selected for the Impact Study, each grantee was required to be:

- Operational (recruiting and matching students to mentors) in Fall 2005 for the first group of grantees and Fall 2006 for the second group, and having mentors meet with students by November of the respective study year.

- Able to over-subscribe or identify excess demand supporting experimental study needs for an un-served control group (i.e., able to provide tangible evidence of a pool of 4th through 8th grade students referred to the mentoring program) of adequate size to support study requirements (a total of at least 40 students per site to yield a minimum of 20 students in each site's experimental (treatment) group and control group).

- Willing and able to cooperate with the data collection and logistical needs of the national evaluation, including random assignment.

A total of 32 unique grantees met these selection criteria and agreed to participate, comprising the final purposive sample. To obtain an adequate sample size of students from the 32 grantees, as described in greater detail in Appendix A, two cohorts of students were recruited for the study. The first phase occurred in the Summer–Fall of 2005, when 21 grantees, and subsequently 1,329 students, were recruited into the study. The second phase took place from Spring–Fall 2006 when 1,244 additional students were recruited from 21 sites.[24]  We collected baseline and follow-up data collection covering approximately one school year for each cohort.

## 2.3 Identification and Random Assignment of Students

When identifying students for the study, grantees used a variety of criteria to determine eligibility, such as grade level or school location. Sites identified appropriate students in a variety of ways, most often asking school staff (such as teachers or counselors from the participating schools) to identify and refer students in need of mentoring. The study sample comprised all eligible students, that is, students who were referred by each program for mentoring, and whose parents signed a consent form to be included in the study. Less than 1 percent of this eligible pool of students was excluded from the study and the requirements of random assignment because the programs considered them to be in extreme need of mentoring services, or because programs were legally bound to deliver services

---

[24]  Of these 32 grantees, 10 provided students and mentors in both rounds of recruitment and random assignment. In other words, 10 of the 21 grantees recruited in 2006 had previously been recruited (and were part of our sample) in 2005, leaving us with 32 unique grantees.  Although 10 grantees provided samples of students and mentors to both study cohorts, students who participated in the first study cohort were purposefully excluded from the second cohort.  In addition, 29 mentors participated in both cohorts of the study, representing 3 percent of the total number of study mentors.

(e.g., cases where students are children of prisoners). Those exceptions were given mentoring, but were excluded from the study to preserve random assignment.

To randomly assign students, the study team randomly ordered each list of students whose parents had consented to their child's participation in the program that had been submitted by individual programs. From these randomly ordered lists, students were sorted into the treatment group by beginning at the top of the list and moving down, selecting as many students as available mentors reported by the program. The lists of students selected for treatment were sent back to programs, so that grantees could begin matching these students.

Those students at the bottom portion of the list whose names were not sent to the program were placed into the control group. These students comprised an ordered wait list for slots opening up if students dropped out or new mentors became available within one month of the date that the program began matching students with mentors. These students were treated as member of the treatment group. In total, approximately 2 percent of students in the control group moved into the treatment group from the wait list.

To preserve random assignment, the study retained all students originally assigned to treatment; that is, program dropouts remained members of the treated group. As described in detail in Section 2.6, the model underlying the random assignment and analysis of students was based on an *Intent-To-Treat* (ITT) framework. In other words, a student's experimental *status* as a treatment or control student, rather than the actual *receipt* of mentoring, served as the measure of treatment. As a result, as students were informed by a grantee that they were allowed to receive services from the program (i.e. that they were treatment students), they were considered to be treated students, regardless of whether they actually received mentoring, or of the nature of the services actually received.

Similarly, wait-list (i.e. control group) students who were not moved to treatment status were considered control group students.[25] Additional details on the identification and random assignment of students can be found in Appendix A.

## 2.4   Statistical Power

The power of a statistical test is the probability that a study design can detect a true difference of a given size (effect size) between the groups under study; that is, the probability that a hypothesis that there is no difference in the groups on the measures (the null hypothesis) can be rejected.  When the groups are treatment and control groups with and without the intervention, this difference is the measured impact of the intervention.  The power to detect as statistically significant any true impacts that do occur in this context is influenced by the size of those effects, the significance level used in the test, and the sample size. A commonly used standard of required power is 0.80; that is, studies commonly seek to detect true impacts of some minimum size 80 percent of the time.  This quantity,

---

[25]   Some control group students may have mistakenly received mentoring from the program as "cross-overs." Note that this group of "cross-overs" only included control group students receiving mentoring from ED-funded Student Mentoring Program grantees. However, control group students were free to obtain mentoring services from other programs in the community. Any such participation was captured in the follow up Student Survey, but did not affect treatment or control group status.

when expressed in standard deviation units (i.e., the impact divided by the standard deviation of the outcome measure) is known as the "minimum detectable effect size."

**Statistical Power for Overall Impacts**

In developing the study design, we strove to select a sample size of students that would be needed to detect overall impacts within the range of effect sizes found previously in the DuBois, et al. (2002) meta-analysis of previous mentoring studies (i.e., ranging from 0.09 s.d, to 0.19 s.d. depending on the outcome domain). The enrolled sample included 42 groups of students from 32 programs, with a total sample size of 2,573 students.  Based on sample size, follow-up survey response rates, the distribution of students across treatment and control groups, model specification, and the explanatory power of included covariates, minimum detectable effect sizes in our study ranged from 0.101 to 0.176 across our set of outcome variables when testing the null hypothesis at a confidence level of alpha = 0.05. Exhibit 2.1 lists the range of minimum detectable effect sizes associated with overall impacts by outcome measure.

**Exhibit 2.1**

**Achieved Statistical Power for Overall Impacts—Minimum Detectable Effect Sizes with Power of 0.80, by Outcome Measure [a]**

| Variable | Sample Size | Minimum Detectable Effect Size [b] |
|---|---|---|
| *Student Survey Outcomes* | | |
| Pro-social Behaviors | 2289 | 0.103 |
| Scholastic Efficacy & School Bonding | 2311 | 0.113 |
| Future Orientation | 2329 | 0.102 |
| Misconduct | 2294 | 0.110 |
| Delinquency | | |
| *Grades* | 1677 | 0.123 |
| Math | 1692 | 0.123 |
| English Language Arts | 1633 | 0.126 |
| Science | 1563 | 0.130 |
| Social Studies | | |
| *Statewide Assessment Tests* | 1840 | 0.117 |
| Math | 1837 | 0.116 |
| Reading/English Language Arts | | |
| *Disciplinary Infractions* | 1847 | 0.131 |
| Misconduct | 1847 | 0.135 |
| Repeated Misconduct | 1847 | 0.132 |
| Delinquency | 1847 | 0.152 |
| Repeated Delinquency | | |
| *Attendance* | 1978 | 0.121 |
| Absenteeism Rate | 1374 | 0.180 |
| Truancy Rate | 1374 | 0.176 |

[a]    For a description of the student outcomes, refer to Section 2.7.

[b]    Expressed as a proportion of a standard deviation.

**Statistical Power for Subgroup Impacts**

The study was designed to detect effect sizes of a given magnitude for the sample as a whole. In contrast, minimum detectable effect sizes are larger for subgroup analyses due to segmentation of the full sample. As a result, the subgroup analyses presented in this report may be underpowered. As seen in Exhibit 2.1, the largest MDE for the student sample as a whole was 0.176, for the Truancy Rate. The smallest was 0.101, for the Pro-social Behaviors scale. Minimum detectable effects were therefore calculated for the largest and smallest subgroups for those two scales with the following results:

1. The smallest subgroup was the group of students who reported delinquent behaviors at baseline. For that subgroup, the MDE ranged from 0.212 (for the Pro-social Behaviors scale) to 0.411 (for the Truancy Rate).

2. The largest subgroup was the group of students with no delinquent behaviors at baseline. For that subgroup, the MDE ranged from 0.117 (for the Pro-social Behaviors scale) to 0.196 (for the Truancy Rate).

We also calculated minimum detectable differences in effect sizes between two subgroups—that is, the smallest difference in true impact between two subgroups with an 80 percent chance of detection with a confidence level (in testing the null) of alpha=0.05. In general, the more evenly divided the sample is between subgroups, the lower the resulting minimum detectable difference in impact between the two populations.

In our analytic sample, the most even division between subgroups was for boys versus girls (47 percent/53 percent split). Minimum detectable difference in impact between boys and girls ranged from 0.204 in effect size units for Pro-social Behaviors to 0.380 for the Truancy Rate.

Conversely, the most unequal split between subgroups was for students reporting delinquent behaviors at baseline versus students with no delinquent behaviors at baseline (25 percent/75 percent split). The minimum detectable difference in impact between students with delinquent behaviors at baseline and students with no delinquent behaviors at baseline ranged from 0.241 in effect size units for Pro-social Behaviors to 0.474 for the Truancy Rate.

Appendix A includes detailed assumptions and formulas for performing the presented power calculations.

## 2.5      Data Describing Grantee and Mentor Characteristics, and Program Implementation

A key component of the study was based on the collection of descriptive data on both program grantee and mentor characteristics as well as on program implementation. This section describes the survey instruments used to measure these characteristics.

**Grantee and Mentor Surveys**

The data collection timeline for gathering information about grantees and mentors is summarized in Exhibit 2.2. Appendix B includes copies of the Grantee and Mentor Survey instruments.

**Exhibit 2.2**

**Data Collection Schedule for Grantees and Mentors**

| Instrument | Timeline | | | |
|---|---|---|---|---|
| | Fall '05 | Spring '06 | Fall '06 | Spring '07 |
| First cohort Grantee Survey | | ✓ | | |
| Representative sample Grantee Survey | | ✓ | | |
| Second cohort Grantee Survey | | | | ✓ |
| First cohort Mentor Survey | | ✓ | | |
| Second cohort Mentor Survey | | | | ✓ |

**Grantee Survey**

The Grantee Survey was developed to collect background information about grantees and program implementation and to measure the extent to which grantees adhere to ED program requirements. This survey collected data on the nature of program experience, program focus, size, budget, staffing, and partners; demographics of the mentors and students; and the amount of training and assistance provided to mentors.

Study sites completed the Grantee Survey during the spring of the year in which they participated in the study (Spring 2006 and/or 2007).

**Mentor Survey**

The Mentor Survey was designed to collect information about mentor demographics and experiences in ED Student Mentoring Programs, and provided further information about program implementation. The Mentor Survey collected data on characteristics such as mentors' relationships with their students, motivation for participation, screening and training, and plans for continued participation.

Mentors were asked to complete two sets of questionnaires. The first set contained questions about the mentors themselves and their general experiences with the Student Mentoring Program (n = 974). The second set included specific questions about mentors' relationships with their assigned students (n = 1,057). Mentors completed student-specific questionnaires for each student with whom they were matched in the Student Mentoring Program.  Mentors were surveyed in Spring 2006 and Spring 2007.  All mentors who were matched with students participating in the study were surveyed (even if they did not actually meet with their students).[26]

---

[26]  A total of 29 mentors were included in both survey administrations because they participated in the Student Mentoring Program with new students during both years of the study. In our analysis, we included data for these mentors from both survey administrations. Although this may seem duplicative, because the mentor is, in effect, the intervention, we wished to consider mentor characteristics and experiences separately for each student paired with that mentor.

## 2.6 Selection of a Random Sample of Comparison Grantees

As discussed in the previous section, data capturing program characteristics and implementation were collected for each of the 32 purposively selected programs. In addition, the Grantee Survey was administered to a random sample of 100 grantees from the pool of 245 eligible programs. The purpose of the random sample was to provide additional descriptive information to ED for program improvement purposes, as well as to determine whether observable characteristics for the purposive sample were comparable to those for the universe of ED program grantees.

For the comparison sample of grantees, the goal was to select a random sample of 100 grantees that would be representative of all ED-funded Student Mentoring Program grantees. Therefore, we stratified programs by auspice (nonprofit/community-based organization, faith-based organization and school district) and by year of funding (2004 or 2005), and then selected a random sample of programs within each stratum. This stratification approach was rooted in the assumptions that: (1) programs operating under different auspices may have recruited different types of mentors, and (2) programs funded in different years may have represented different levels of experience or stability at the time of the Grantee Survey.

Using stratification in the selection of grantees was beneficial in two ways: (1) it guarded against extreme cases where non-representative draws may have occurred by chance,[27] and (2) it could also reduce the variance of overall estimates of program attributes by eliminating variation across strata. To minimize the variance, we stratified the sample in proportion to the number of grantees in each stratum among the 245 funded eligible programs.

Although the stratification and random selection of these 100 grantees occurred concurrently with grantee recruitment for the purposive sample for the impact evaluation, these processes were completely independent. That is, the purposive selection of a grantee for the evaluation had no bearing on that grantee's selection for the larger random sample of representative sites, and vice versa. Consequently these simultaneous and nonexclusive selection procedures resulted in an overlap of 12 grantees between these two groups. Further details on the selection of grantees for the comparison sample can be found in Appendix A.

## 2.7 Measuring Student Outcomes

ED's Student Mentoring Program is intended to provide students at risk for poor academic and/or behavioral outcomes with a mentor who serves as a role model in regular contact with the student, builds a positive relationship, and participates in the activities programs support. We based the evaluation outcome measures on the intended outcomes as stated in the legislation authorizing the program: better school or academic performance and engagement (grades, academic test scores, scholastic efficacy and bonding to school, attendance), improved desired behaviors (interpersonal relationships with adults, personal responsibility, and community involvement), and reduced delinquent behaviors (misconduct in school and at home, truancy, drug, alcohol and tobacco use, gang involvement). Since school based mentoring programs operate for at least one school year, our

---

[27] This was especially important because some groups of grantees (e.g., faith-based organizations) made up a relatively small proportion of the relevant universe, and could have potentially been left out of a simple random sample drawn without stratification.

design collects data on these outcomes for the randomly assigned students from the time of their random assignment in the fall of the school year to the following spring.

As described below, some outcome data come from the students themselves through surveys administered in the fall and in the spring of the program year. These data include self-reports of attitudes toward school, relationships with parents and other adults and self reported misconduct or delinquency. Other data come from school records (grades, performance tests, disciplinary actions, attendance, and truancy) abstracted for all study sample students reflecting the year before the study and at the end of the school year in which the study took place. Exhibit 2.3 presents the data collection timeline for student data.

**Exhibit 2.3**

**Data Collection Schedule for Students**

|  | Timeline | | | |
|---|---|---|---|---|
| **Instrument** | **Fall '05** | **Spring '06** | **Fall '06** | **Spring '07** |
| First cohort Student Survey | ✓ | ✓ | | |
| Second cohort Student Survey | | | ✓ | ✓ |
| First cohort student record abstraction | | ✓ | | |
| Second cohort student record abstraction | | | | ✓ |

The instrumentation for student data collection was guided by the objectives of the student mentoring grant program outlined in the authorizing NCLB legislation. We briefly describe below the two sources of impact data (Student Surveys and school records) and address the construction of outcome measures from these sources in more detailed sections that follow.

**Student Surveys**

Student Surveys were one source of data used to analyze program impacts for students, collected for individual students in both the treatment and control groups. The surveys were collected initially in the fall of the relevant program year and then again in the spring for each cohort of students. The goal in developing the Student Survey was to reflect domains directly related to the legislative intent of the program as outlined in the program logic model in Exhibit 1.1. The Student Survey drew on reliable measures used in other surveys, adapted for the needs of the present study.

All randomly assigned students completed a baseline survey, preferably before random assignment.[28] For the follow-up survey, students in both treatment and control groups were notified and assembled for the survey at the same time. All of the Student Surveys were group-administered in school-based settings. Appendix B contains copies of both the fall and spring Student Surveys.

---

[28]    In some cases, for example when a student was absent from the initial survey administration, the survey was administered post-random assignment. In almost all cases, however, the survey was administered prior to receipt of mentoring on the part of treatment group students.

**School Records**

School records for individual students were also used to measure program impacts on student outcomes. Where possible, school records were collected for each student in the spring of the year in which he or she participated in the study, as well as for the preceding school year. These records included data on:

- Student grade level and demographics;
- Class performance in math, English language arts, science and social studies;
- Performance on statewide assessment tests in math and reading/English language arts (ELA);[29]
- Disciplinary infractions; and
- Truancy and absence from school.

**Construction of Student Outcomes: Student Surveys**

In this section, we outline our approach to constructing student outcome measures from Student Surveys. Exhibit 2.4 summarizes the sources and measures for the student outcomes. The measures included in the Student Survey instrument were chosen to reflect the specific goals of the mentoring program as laid out in the logic model (Exhibit 1.1) in the following three impact domains:

1. Improved Interpersonal Relationships, Personal Responsibility, and Community Involvement
2. Improved Academic Outcomes and Participation
3. Reduced Juvenile Delinquency/Participation in Harmful Activities

The Student Survey incorporated a number of measures mapping to each impact domain, as detailed in Appendix C, Exhibit C.1. As shown in this exhibit, most of these measures were derived from existing scales with adequate levels of reliability and validity previously established in prior research. However, because some of these scales were developed on older populations of students than those represented by our study sample, and/or contained large numbers of individual questions or subscales of varying relevance to the impact domains of interest, most were altered in some way prior to their inclusion in the survey instrument. For our study sample, many of these Student Survey scales did not meet standard minimal criteria for internal reliability.

---

[29] School districts and states differed in terms of how this construct was measured. A total of 31 sites reported state proficiency test information in either reading or English language arts. Of these, 24 sites reported assessment scores in reading, 6 sites reported scores in English language arts, while 1 site reported scores in reading one year, and in English language arts the next.

**Exhibit 2.4**

**Sources and Measures for Student Outcomes**

| Student Outcome | Source | Measure |
|---|---|---|
| *Self-Reported Attitudes and Behaviors* | | |
| Pro-social Behaviors | Student Survey | Sum of standardized scale items |
| Future Orientation | Student Survey | Sum of standardized scale items |
| Misconduct | Student Survey | Sum of standardized scale items |
| Delinquency | Student Survey | Sum of standardized scale items |
| Scholastic Efficacy and School Bonding | Student Survey | Sum of standardized scale items |
| *School Performance: Assessment Tests/Grades* | | |
| Grades: English Language Arts, Math, Science, Social Studies [a] | School Records | A – F (1-5) |
| Statewide Assessment Tests: Reading/ELA and Math | School Records | Student meets state proficiency requirements threshold score (0,1) |
| *Official Disciplinary Infractions/School Attendance* | | |
| Truancy | School Records | Total number of unexcused absences from school as a proportion of total school days required |
| Attendance | School Records | Total number of absences from school as a proportion of total school days required |
| Misconduct (Any, Repeated)[b] | School Records | Disciplinary actions for Misconduct recorded by school for student |
| Delinquency (Any, Repeated)[b] | School Records | Disciplinary actions for Delinquency recorded by school for student |

[a]  Not all schools reported science and social studies grades at the elementary level.

[b]  See Appendix C, Exhibit C.6 for a listing of infractions that make up these categories.

To correct for potential threats to internal validity in developing final outcome measures from Student Survey data, we performed an exploratory factor analysis with Promax rotation to refine and confirm scale construction for our sample. Appendix C describes in detail the factor analytic steps taken in developing the final Student Survey outcome measures, which maximize internal reliability while preserving a logical mapping to the impact domains of interest. The five Student Survey outcome indices created by the factor analysis are:

1. Pro-social Behaviors (Interpersonal Relationships, Personal Responsibility, and Community Involvement)
2. Future Orientation;
3. Misconduct;
4. Delinquency; and
5. Scholastic Efficacy and School Bonding.

Appendix C, Exhibit C.2 provides reliability coefficients and lists the individual Student Survey question items for each of these five outcome measures.[30]

---

[30]  See Appendix D for impact estimates based on scales as they appeared in the Student Survey. We caution the reader to note that, although these measures are more directly comparable to scales used in prior

All items comprising each scale except the Delinquency measure were measured based on the same four-point Likert scale. Four items in the Delinquency measure were scored using different metrics. The drug, alcohol, and tobacco items were answered in terms of frequency of use in six ordinal categories, from "never used" and "none" to "10 or more times" in the prior 30-day period. The gang involvement items were dichotomous (i.e., currently in a gang or not). All of these items were standardized to the mean and standard deviation of the two Likert-scale Delinquency measures.

**Construction of School Records Outcomes: Grades and Statewide Assessments**

Data on grades and statewide assessment scores from school records were used to produce outcome measures intended to assess impacts on academic achievement. This subsection details construction of final academic outcome measures from the school record data.

**Grades.** Grades were collected for each student for the school year prior to the study year and for the year (spring semester) of the study. Schools differed in what subjects were graded, depending on the school or district location and the grade level of the student. Not all students across sites had grades recorded in each subject category. We created outcome measures of grades in the following four subject areas, which were available for most students and schools in our sample: English language arts, math, social studies, and science.

Many school districts (and even schools within districts) adopt their own conventions in measuring grades, so we created a common measure of grades, allowing us to make relative comparisons of the magnitudes of impacts across sites. The vast majority of school districts assigned grades either on a 0–100 numeric scale or on an "A–F" letter scale.

Since conventional numerical grading systems generally use similar number-to-letter-grade translations, we were able to establish a reasonable translation of numeric scores to letter ratings to create a categorical grade measure comparable across sites. Appendix C, Exhibit C.4 shows how numeric ranges, letter grades, and other grading schemes were transformed to a categorical 1–5 performance level, with 1 representing the lowest performance level and 5 the highest. Grades were then further standardized across each site by dividing each score by the within-site standard deviation.

It is unlikely that this approach reconciled all differences in grading across sites in the evaluation, but a complete reconciliation is unnecessary given our use of a fixed effect model, as described in greater detail in Section 2.6. Specifically, the fixed effect model will capture all systematic differences in grading across schools and sites. The treatment effect for a fixed effect model is an estimate of how the treatment shifted the grades for the treated and untreated *within* a program, so the estimated treatment effect for a site will be estimated consistently even if grades are measured inconsistently across sites.[31]

---

studies, only two of the eight scales comprised of multiple items met our minimal criteria for internal reliability.

[31]  A model with program-level fixed effects will not capture differences across schools *within* each site. However, because grading procedures for schools within each site are likely to be more alike than grading procedures for schools across all sites, we expect that the site-level fixed effect specification will sufficiently capture differences.

**Statewide Performance Assessments.**  Schools also provided statewide performance test scores for students in reading/ELA and math. Not surprisingly, testing protocols varied widely across states. In constructing our reading/ELA and math proficiency outcome measures, we obtained proficiency thresholds for each state's performance tests and coded each sample student's scores into a dichotomous score (proficient, non-proficient) based on those thresholds. Baseline proficiency measures were based on test results for the spring prior to the study year, and post-treatment proficiency measures on test results for the spring of the study year. Appendix C, Exhibit C.3 summarizes assessment coding rules by site.

**Construction of School Records Outcome: Disciplinary Infractions and School Attendance**

We collected data on disciplinary infractions and attendance from each student's school record and coded them into standardized measures of delinquency, misconduct, attendance, and truancy. The remainder of this subsection details the steps taken in constructing these standardized outcome measures.

**Delinquency and Misconduct.** School records contained a wide variety of terms to describe the range of reported infractions. In order to derive infraction measures that were consistent across sites, we first sorted all infractions reported into seven categories (harassment, general non-compliance, property offenses, drug related infractions, truancy, violence, and other), and totaled the number of reported infractions within each category for each student in our sample. These infraction categories were then grouped into two broader classifications: misconduct, representing less serious infractions (non-compliance, harassment, and property offenses), and delinquency, representing more serious offenses (violence, drug related infractions and truancy). Appendix C, Exhibit C.5 lists the individual infraction types contained in each categorization. For each of the two broad infraction categories, we constructed outcome measures to indicate if students had records of repeated infractions: any misconduct, repeated misconduct, any delinquency, and repeated delinquency.[32]

**Attendance and Truancy.** We examined school records to derive measures of attendance, which incorporated both excused and unexcused absences, and truancy, which incorporated unexcused absences only. Attendance was defined as the proportion of days that each student in the study attended relative to the number of days in the district's school year, regardless of whether the absence was excused or unexcused. Truancy was defined as the proportion of unexcused absences relative to the total number of days in the district's school year.

## 2.8   Response Rates

Exhibit 2.5 shows response rates for the Grantee, Mentor and Student Surveys,[33] as well as for outcome measures obtained from student records. In general, survey response rates were quite high, while response rates for student record data items were somewhat lower. Some school districts, for

---

[32]   Given the highly skewed distributions of the infraction measures, where most students had no incidents and relatively few students had a high number of incidents, we were not able to establish ordered categories for these measures that would meet the assumptions of standard statistical tests for ordinal measures.

[33]   As described in greater detail in Section 2.9, we imputed Student Survey responses for some individual items where appropriate. However, the Student Survey response rates reported in Exhibit 2.5 reflect rates calculated prior to the implementation of the imputation procedure.

example, were not able or were unwilling to provide data. We present sensitivity analyses examining the influence of missing data on our estimated impacts in Appendix E.

**Exhibit 2.5**

**Response Rates by Survey or Data Type**

| | Number of Surveys Distributed | Number of Surveys Returned | Response Rate |
|---|---|---|---|
| *Student Survey* | | | |
| Baseline (Fall 05 & Fall 06) | 2573 | 2529 | 98.3% |
| Treatment | 1272 | 1244 | 97.8% |
| Control | 1301 | 1285 | 98.8% |
| Post-Treatment (Spring 06 & Spring 07) | 2573 | 2377 | 92.4% |
| Treatment | 1272 | 1174 | 92.3% |
| Control | 1301 | 1203 | 92.5% |
| *Mentor Survey* | 1138 | 974 | 85.6% |
| *Grantee Survey* | | | |
| Random Sample | 100[a] | 100 | 100.0% |
| Impact Study Sites | 32[b] | 32 | 100.0% |
| *Student Record Abstraction* | **Number of Students** | **Number of Records** | **Response Rate** |
| Grades | | | |
| Year 1 (04-05 & 05-06) | 2573 | 1842 | 71.6% |
| Year 2 (05-06 & 06-07) | 2573 | 2099 | 81.6% |
| Statewide Assessments | | | |
| Year 1 (04-05 & 05-06) | 2573 | 2062 | 80.1% |
| Year 2 (05-06 & 06-07) | 2573 | 2133 | 82.9% |
| Disciplinary Infractions | | | |
| Year 1 (04-05 & 05-06) | 2573 | 1933 | 75.1% |
| Year 2 (05-06 & 06-07) | 2573 | 2045 | 79.5% |
| Attendance | | | |
| Year 1 (04-05 & 05-06) | 2573 | 2289 | 89.0% |
| Year 2 (05-06 & 06-07) | 2573 | 2450 | 95.2% |

[a]   Includes 12 surveys from Impact Study sites.

[b]   Ten grantees from Cohort 1 were also surveyed for a second time.

The response rate for the Grantee Survey for both groups of grantees to which it was administered was 100 percent. Although some grantees did not complete all questions and some indicated that they did not collect some of the data requested, the response rate for most questions was 85 percent or higher. The response rate for the Mentor Survey was 86 percent.

Student Survey response rates in both cohorts were similar for treatment and control groups, ranging from 92 percent for the spring sample to 98 percent for the fall sample. Response rates were somewhat lower for year 2 school record abstraction data items, ranging from 80-95 percent for grades, state assessments, disciplinary actions, and attendance data. In general, response rates for school record data were higher in year 2 of the study, when records for both years were collected.

## 2.9    Impact Analysis Methods and Procedures

This section describes our approach to analyzing the impact of the ED Student Mentoring Program on eligible students.

**Estimation of Overall Impact**

Our statistical model is based on a multi-site randomized trial design. Because the random assignment of students to treatment and control groups occurred at the individual site level, we estimated the overall program impact by **averaging separately derived site-level impacts**, that is, we first estimated an impact within each site and derived an average impact across the selected sites.  The statistical model was thus based on a two-stage estimation strategy, whereby in the first stage each outcome variable was modeled within each site yielding 42 site-specific impacts.  These impacts were then averaged in a subsequent stage to yield an overall impact estimate for each outcome variable.  In addition, the statistical model was based on an *Intent-to-Treat* framework testing the effect of making mentoring available to eligible students. That is, the experimental design measures the impact of School-based Mentoring on all students assigned to the treatment group, whether or not they actually received mentoring services.

Within a randomized experimental design, valid impact estimates can be based on simple comparison of mean outcomes between the entire treatment and control groups when the probability of selection into the treatment group is the same across program sites. In the impact evaluation, however, the probability of being assigned to treatment varied across program sites and across points in time and/or schools within sites, necessitating the use of observation weights. Our statistical model was further enhanced through the use of site-level fixed effects, which controls for differences in mean outcomes across program sites, and through the inclusion of baseline covariates, which increases model precision.

*Observation Weighting*
In a number of sites, programs were obliged to select participants in multiple batches at different points in time or for different schools within each site, using different treatment/control group assignment probabilities in different batches. Since population characteristics may have varied for the different batches, data across batches had to be balanced between the treatment group and the control group.

This was accomplished by weighting each observation by the inverse probability of selection into the relevant group. This weighting strategy preserved the balance between the treatment and control groups in terms of the mix of students from different randomization batches, and preserved random assignment as the basis for estimating the treatment effect, thus eliminating potential bias arising from unequal random assignment probabilities across time or schools within a site .  In other words, the use of observation weighting adjusts for the clustering of students by school.  This, in effect, accounts for any school-level effects in the impact estimation.

Specifically, the observation weight ($\omega_{ijt}$) was defined as:

$$[2] \quad \omega_{ijt} = \frac{1}{P_{ijt}} \text{ (where } T_{ijt} = 1\text{), and}$$

$$[3] \quad \omega_{ijt} = \frac{1}{1 - P_{ijt}} \text{ (where } T_{ijt} = 0\text{)}$$

where,

$P_{ijt}$ = the achieved probability of being selected into the treatment group for student $i$ in site $j$ at time $t$ (i.e. at the time that random assignment was conducted for student $i$ for his or her batch).

To simplify subsequent notation, we represent the weighting term ($\omega_{ij}$) and the treatment indicator term ($T_{ij}$) without the $t$ subscript in the remainder of the text.

*Model Specification*

Weighted Least Squares (WLS) regression was used to estimate the model incorporating the observation weights to determine the overall impact of student mentoring:

$$[4] \quad \omega_{ij} Y_{ij} = \sum_{j=1}^{J} \omega_{ij} \, \beta_{1j} T_{ij} \, S_j + \sum_{j=1}^{J} \omega_{ij} \beta_{2j} S_j + \omega_{ij} \beta_3 \, X_{ij} + \omega_{ij} \, \varepsilon_{ij} \, ,$$

where,

$Y_{ij}$ is the outcome of interest $Y$ for student $i$ in site $j$,

$T_{ij}$ is the treatment indicator for student $i$ in site $j$ ($T_{ij} = 1$ if student $i$ is assigned to the treatment group; $T_{ij} = 0$ otherwise),

$S_j$ is a site indicator equal to 1 for students randomized at site $j$ and to 0 otherwise ($j = 1...J$),

$\beta_{1j}$ is the estimated average ITT treatment effect for site $j$,

$\beta_{2j}$ is the program fixed effect at site $j$ (i.e., the average untreated outcome level of a student at site $j$),

$X_{ij}$ is a vector of student characteristics measured for each student $i$ in site $j$,

$\beta_3$ represents the vector of coefficients indicating how student characteristics affect student outcomes, and

$\varepsilon_{ij}$ represents a random error term for student $i$ in site $j$, independent and identically distributed across students.

The estimated variance[34] of the weighted least-squares impact estimate $\hat{\beta}_{1j}$ , for $K$ total parameters estimated in equation [4], is

$$[5] \quad Var(\hat{\beta}_{1j}) = \frac{\hat{\sigma}_j^{\,2}}{\sum_{i=1}^{N}(T_{ij} - \bar{T}_{ij})^2} \quad \text{for each given site } j = [1...J],$$

where,

$$[6] \quad \hat{\sigma}_j^{\,2} = \sum_{i=1}^{N} \frac{(Y_{ij} - \hat{\beta}_{1j}T_{ij} - \hat{\beta}_{2j} - \hat{\beta}_3 X_{ij})^2}{N - K + 1} \cdot {}^{35}$$

### Site-level Fixed Effects

We use a fixed effects model in this analysis, that is, the model does not account for variation across study sites. This approach is taken because sites were not selected to be a random sample.[36] The introduction of $J$ site indicator variables ($S_j$) and their corresponding estimates ($\beta_{2j}$) in equation [6] implies this fixed effect model. The $J$ fixed effects ($\beta_{2j}$) capture variation across the average student outcome level (for both treatment and control students) for individual sites.[37] For example, if treatment and control students from site "A" had worse grades on average than treatment and control students from site "B" after adjusting for baseline student characteristics $X_{ij}$, the fixed effect estimates controls for these average differences, making the two sites more comparable.

The fixed effects specification similarly accounts for any average differences across sites arising from the construction of outcome measures based on the diverse information contained in school records data, as described in Section 2.5. For example, site "C" may have employed higher standards for judging its students "proficient" in math based on assessment tests than site "D." Provided that the mean difference in standards remained consistent between the pre- and post-treatment periods, the inclusion of fixed effect indicators in this model specification controls for these differences.

### Student Characteristics

The precision of the impact estimates was improved by controlling for the baseline characteristics of students that are related to outcomes and not explained by treatment. The inclusion of a vector of

---

[34] Appendix E presents impact results incorporating heteroscedastic-robust standard errors. This alternative approach did not result in any changes in the statistical significance of the estimates.

[35] Note that these are the standard equations for the error variance in an ordinary least squares regression, and do not include sample weights. A derivation of these equations in this context is available from the authors on request, but is omitted here for the sake of brevity.

[36] We use the term "fixed-effects" within the dual perspectives of sampling and statistical inference. Because student mentoring programs were chosen purposively, not randomly into the study, results cannot be generalized to the full universe of programs. This model is therefore, appropriate, given our level of inference does not extend beyond our study sample of purposively chosen programs.

[37] The fixed-effect site-level dummy variables account for any clustering effects that may occur within a site attributable to site identity.

observable baseline student-level characteristics $X_{ij}$ in our model achieved higher levels of statistical power by removing these controlled sources of variation from the error term in the impact model.

In addition, even under random assignment, statistically significant baseline inter-group variations in characteristics may result from random chance. Including baseline characteristics in our model also helped to adjust for chance differences occurring despite random assignment, as well as differences in individual student characteristics across the overall sample that may have affected outcomes. Included covariates are:

- Baseline value of outcome measure being predicted;
- Age;
- Gender;
- School lunch eligibility status;
- Race/Ethnicity; and
- Family structure (two-parent households versus all others).

These covariates were chosen on the basis of their theoretical importance in explaining variation in outcome measures of interest. Although not all covariates were necessarily thought to influence variation in all individual outcomes, we included the same set of covariates in the model for each outcome to maintain consistency in this approach. Statistical tests (F ratio tests) indicated that these covariates were statistically significant predictors of outcomes, so they are included in the models to improve statistical power.

***Pooled Impact Estimates***
The specification described in equation [4], when estimated in a single step by WLS regression, led to *J* estimates of the treatment effect ($\beta_{1j}$), one for each of the *J* sites. To obtain a composite estimate of the treatment effect, we had to compute a single aggregate estimate from these *J* estimates, representing an average effect across sites. However, the average effect will vary depending on the relative weight attached to each site-level impact estimate. The weighting methodology employed depends on the precise research question to be answered, which can be conceptualized in several ways:

- What is the average effect per mentoring program?
- What is the average effect per student assigned to mentoring?
- What is the average effect per student eligible to receive mentoring?

The first question attaches equal importance to each mentoring program, and therefore implies that each site should be given equal weight when calculating the average effect. The latter two questions attach greater importance to larger mentoring programs, and therefore imply that sites should be weighted proportionally to the number of students assigned to the treatment group or to the number of students in the study, respectively.

Based on the substantial variation in site sample sizes (the largest site has three times as many students as the smallest), we chose to weight sites proportionally to the total number of treatment and control students at each site. In calculating the average treatment estimate, each site was weighted by the total site sample size divided by the total number of students in the study, such that all weights sum to one. The final reported impact estimates could therefore be interpreted as the average

treatment effect per student eligible to receive mentoring, although the average is to be interpreted within an Intent-to-Treat framework.[38]

**Imputing Missing Data**

In cases with missing data for covariates, we employed a mean substitution method, whereby the site-level mean covariate value was substituted for any student with missing data for that value. Mean substitution is a conservative method of data imputation because it reduces the variance in the covariate, thereby decreasing the ability of the covariate to explain differences in the outcome measure. Increased measurement error resulting from imputation will tend to inflate the estimated slope coefficients on the covariates with missing data; however, it will not alter the coefficient on the treatment status variable, which is never missing in our sample. For the purposes of this evaluation, it is therefore an acceptable means of including all cases with outcome data in our regression analyses without introducing bias into treatment effect estimates.

**Estimation of Subgroup Effects**

Existing literature suggests that the impact of student mentoring programs may differ across student subgroups such as boys and girls or age categories (DuBois et al., 2002). Therefore, we conducted subgroup analyses to determine whether the student mentoring program differentially affected certain subsets of students. We analyzed subgroups differing by (1) gender, (2) age (students 12 or older versus students less than 12 years old), (3) family structure (students from two-parent families versus students from other types of families), (4) presence of self-reported delinquent behaviors at baseline (theft, possession of a weapon, drug use, alcohol use, or gang activity), and (5) academic non-proficiency (in math and/or reading/ELA) at baseline.

To estimate effects for these subgroups, we first divided the entire sample according to the characteristic of interest (e.g., one sample entirely comprised of boys, the other sample comprised of girls), to preserve randomization. We then estimated the treatment effect for each subsample by site using an analogous WLS regression specification to that described for the full sample above. Site-level impacts were averaged using weights proportional to the size of the treatment and control group in each subgroup for each site. This produced separate aggregate impact estimates for each subgroup; for example, one impact estimate for boys, and one for girls.

Finally, a t-test was used to test for statistically significant differences in the magnitude of impacts between the two paired subgroups. This tested not whether impacts were present for boys or for girls, but whether impacts on boys and girls statistically significantly differed.

**Multiple Comparisons**

Multiple comparisons in our study arise in two ways: (1) in the assessment of treatment effects on multiple outcome measures, and (2) in comparing multiple outcomes across various subgroups within our sample. When performing multiple hypothesis tests, the likelihood of finding a "statistically significant" effect increases with the number of hypotheses being tested. For example, if we were to

---

[38]      Appendix E presents sensitivity analyses, including a comparison of impact estimates under alternative weighting schemes. In general, these alternative weighting approaches did not significantly alter our findings.

perform twenty tests, we would expect (on average) to find one statistically significant difference at the 0.05 level simply by chance. Without accounting for multiple comparisons, we might incorrectly claim that student mentoring programs statistically significantly influenced some outcomes even when they did not.

### *The Benjamini-Hochberg Correction*

The false discovery rate (FDR) is the expected proportion of all rejected null hypotheses that are false discoveries, or results caused by chance alone. To correct for the use of multiple comparisons, we adjusted for the FDR using a method developed by Benjamini and Hochberg (1995), hereafter referred to as the BH correction. The BH correction establishes an upper bound on the FDR for the purposes of hypothesis testing; the true FDR will generally be lower than the rate it imposes. Compared to other multiple-comparison correction procedures, the BH correction generally enjoys greater statistical power while also being more robust to the variation in comparisons conducted (Williams, Jones and Tukey, 1999).

In the BH correction, the following four-step procedure was carried out to adjust for the FDR:

1. Conduct N separate t-tests for each outcome to test the null hypothesis that no differences exist, each at level $\alpha$.

2. Order the observed p-values from smallest to largest, where $p_{(1)} \leq p_{(2)} \leq p_{(3)} ... \leq p_{(N)}$.

3. Define *k* as the maximum *j* than satisfies the condition: $p_{(j)} \leq c_{(j)} = \frac{j}{N}\alpha$.

4. If *k* does not exist, then no $H_0$ will be rejected; if *k* exists, then reject $H_{(0)j}$, where *j*=1, 2,…, *k*.

The BH correction does not provide an adjusted p-value for each t-test. Instead, it yields a series of critical p-values (the "$c_{(j)}$" values calculated in step 3 above) to which each observed p-value is compared to determine significance. The impact analysis results tables in Chapter 4 provide the BH-corrected critical value in a separate column; unadjusted p-values less than the BH-corrected critical value represent results that are statistically significant after adjusting for the false discovery rate.

### *Defining Families of Comparisons*

In applying the BH correction (or any other procedure intended to adjust for multiple comparisons), one must first determine the "families" of hypothesis tests across which we wish to adjust.

For the purposes of this study, we have defined families of comparisons to coincide with the three impact domains of interest. This approach is consistent with the program logic model and the overall intent of the study. Similarly, for our subgroup analyses, families of comparisons are defined for each impact domain within each individual subgroup, as opposed to across all five sets of subgroups simultaneously.

# Chapter 3: Characteristics of Grantees and Mentors; and Program Delivery

This chapter of the report provides descriptive information about the 32 grantees that participated in the Impact Study as well as the extent to which they were similar to the universe of grantees in ED's Student Mentoring Program. The chapter also provides background information on the mentors who participated in the study.

The NCLB legislation authorizing the Student Mentoring Program provided general guidelines for how grantees were to use their funds in recruiting students and mentors, and overseeing mentor activities with students. As previously outlined in Chapter 1, grantees were instructed to adhere to the requirements of the legislation as well as follow the recommendations provided by the program office. In addition to setting the absolute priority of addressing the academic and social needs of children with the greatest need, OSFDS, in their grant solicitation, also outlined a number of strategies underlying well-designed and effective school-based mentoring programs:

- Measurable program goals;
- Identification of students to be mentored and individuals to serve as mentors;
- Defined expectations for frequency of student/mentor contact, parameters of the student/mentor relationship, and criteria for matching students with their respective mentors;
- Screening of all potential mentors including background checks;
- Training and support for mentors and program staff on an ongoing basis;
- activities for mentors and students; and
- Established procedures for supervising and monitoring of mentoring relationships.

The findings in this chapter are organized to yield a comprehensive picture of program characteristics, including type of organization, prior experience running mentoring programs, and characteristics of students served by the program. In addition, findings are reported on characteristics of the mentors matched with students in the Impact Study sample. The findings also include information on program delivery including training and support provided to mentors, process of matching students to mentors, frequency and duration of student/mentor meetings, activities engaged in by mentors and students, and perceptions of the mentoring relationship from the perspective of both mentors and students. Finally, the chapter presents data on the study treatment contrast, specifically in terms of the extent of other mentoring activities in the community provided to students in both groups.

# 3.1   Overview of Findings

The key descriptive findings presented in this chapter include the following:

***Impact Study Grantee Characteristics:***
- The majority of programs in the Impact Study (66 percent) were operated by nonprofit/community-based organizations or faith-based organizations.

- Among Impact Study sites, experience running school-based mentoring programs ranged from less than 1 year to 35 years.

- Impact Study sites served an average of 217 students (per site) during the 2005-2006 and 2006-2007 school years, with a third (34 percent) serving 250 or more.

- Ninety percent of Impact Study grantees reported being extremely focused on academics, including academic engagement and achievement.  In contrast, thirty-eight percent reported being extremely focused on risk avoidance.

- On average, 86 percent of the students served by the programs in the Impact Study were in grades 4–8; 77 percent were non-white; and 57 percent were female.

- The most common student risk factor reported by Impact Study grantees was the lack of positive adult role models in students' lives, cited by three-fourths (75 percent) of the grantees.

- Among the statistically significant differences between the grantees participating in the Impact Study and a representative sample of randomly selected programs were that grantees in the Impact Study had more experience running mentoring programs, served larger numbers of students, had larger budgets devoted to school-based mentoring activities, and were more likely to be operated by school districts.[39]

***Mentor Characteristics:***
- Seventy-six percent of program mentors in Impact Study sites reported having had some or a lot of prior contact with students in grades 4–8.

- Seventy-two percent of these mentors were female; approximately two-thirds (66 percent) of the mentors were white.

- Approximately one-half (49 percent) of the mentors were employed; 39 percent were in school full-time, primarily in college.

- Eighty-two percent of the mentors had completed at least some college or other form of post-secondary training.

***Program Delivery:***
- Fifty-three percent of the mentors and students were matched in terms of race and ethnicity, and approximately four-fifths (81 percent) were of the same gender.

- Fourteen percent of the students in the treatment were never matched with a mentor during the school year and another 3 percent were matched but never met with their

---

[39]   Other differences are discussed in the chapter text and accompanying exhibits.

mentors.  Of those mentors who met with their students, 87 percent reported meeting with their students on a one-to-one basis.

- The average time between the start of the school year and the date students were matched with a mentor was 81 days .

- Eighteen percent of mentors were 18 years of age or younger and an additional 23 percent were college-age, and 31 percent of all mentors reported previous experience mentoring.

- Mentors averaged 4.4 meetings per month with their student(s) and reported meeting with their student(s), on average, for 5.8 months by June of the relevant school year.[40]

- Discussing students' relationships with others, including parents, peers, teachers and other adults in authority was the most common activity undertaken by mentors and their students.  Fifty-two percent of mentors meeting with students reported discussing students' relationships either "most of the time" or "almost always."  Discussing risk behaviors was the least common with 23 percent discussing this topic "most of the time" or "almost always."  Forty-three percent of mentors reported working on academics with their students either "most of the time" or "almost always."

- Mentors reported a variety of supports being available from grantees.  Ninety-four percent of mentors reported having access to some kind of ongoing supports (such as supervision or access to social workers) and 96 percent received some sort of pre-match training or orientation.

- The overwhelming majority of students and mentors had positive feelings about their mentoring relationships.

*Treatment Contrast:*

Eighty-six percent of treatment group students reported receiving mentoring services over the past school year, primarily from ED Student Mentoring programs, compared to 35 percent of the control group students. (Three percent of students in the control group received mentoring from the grantees in the study). This difference was statistically significant.  However, the mentoring reported by students in the treatment group appeared to be more intensive: 85 percent of students in the treatment group who actually received mentoring met with their mentors at least twice a month, compared to approximately two-thirds (66 percent) of the mentored control group students, a statistically significant difference.

*Context of Findings:*

Grantees reported providing mentoring services consistent with what was outlined in the legislation and what is recommended under guidelines from the program office and the mentoring field. Grantees indicated that they were for the most part serving the appropriate students in terms of age (i.e., on average, sites indicated that 86 percent of their students served were in grades 4-8) and risk factors as specified in the legislation (e.g., three-quarters of grantees indicated that having few or no positive adult role models was one of the top three perceived risk factors for students in their mentoring programs), and targeting the academic and social needs of students as prioritized by OSFDS (i.e., 91 and 84 percent of grantees reported being extremely focused on academics and student self-esteem,

---

[40]    In cases where mentors planned to meet with their students beyond June of that school year, they were instructed to check June as the month they expected to stop mentoring.

respectively). In addition, 87 percent of mentors reported meeting on a one-to-one basis with their students, averaging 4.4 meetings per month with meetings lasting 1.1 hours on average. Mentors also reported having available a variety of supports from their programs, ranging from pre-match training (96 percent of mentors) to supervised meetings (51 percent of mentors). These findings are all consistent with the guidelines provided by the program office.

Although both mentors and students reported having positive feelings about their mentoring relationships (e.g., 94 percent of mentors reported enjoying the time with their student(s) and 89 percent of students reported that they could trust their mentor), the length of their relationships was attenuated by the late start-up experienced by many programs. Finally, the fact that 35 percent of the control group students reported receiving mentoring either from the program and/or elsewhere in the community coupled with the fact that not all treatment group students met with a mentor may have led to some dilution of the anticipated treatment effect.

## 3.2    Characteristics of the Grantees in the Evaluation

As shown in Exhibit 3.1, grantees varied in terms of organization type, experience operating mentoring programs and the number of students served. Although all grantees operated school-based mentoring programs, not all grantees were schools or school districts. Thirty-four percent of the Impact Study site grantees were schools or school districts, while approximately two-thirds (66 percent) of the Impact Study sites were nonprofit/community-based organizations or faith-based organizations. Survey results from our representative sample of grantees were similarly varied, although a higher percentage of these grantees were nonprofit/community-based or faith-based organizations (71 percent) while fewer of them were schools or school districts (29 percent). These differences were statistically significant at the .05 confidence level.[41]

Grantees ranged in their experience running school-based mentoring programs from less than one year to thirty-five years. Impact Study sites had more experience, with an average of 6.1 years experience running school-based mentoring programs compared to 4.9 years for sites from the representative sample of grantees, a statistically significant difference. When we consider the percentages of sites from each group with various levels of experience, 25 percent of sites from the representative sample of grantees had less than two years of experience running school-based mentoring programs, compared to 6 percent of sites in the Impact Study, and this difference was statistically significant.[42] Similarly, 40 percent of the Impact Study sites had six years or more of experience, compared to 32 percent of the sites from the representative sample (and, again, this difference was statistically significant).

---

[41]    All statistically significant differences are reported at the .05 confidence level. Appendix A explains how we determined the statistical significance of differences between grantees in the Impact Study and grantees in the representative sample.

[42]    In addition to comparing the average years of experience operating school-based mentoring programs for grantees in the Impact Study and grantees in the representative sample, we determined the percentage of grantees in each group that had relatively little experience running school-based mentoring programs (fewer than two years) and the percentage that had relatively significant experience running school-based mentoring programs (more than six years). These data are not included in Exhibit 3.1.

As Exhibit 3.1 illustrates, student mentoring programs varied a great deal both in the numbers of students they served and in their school-based mentoring budgets. Annual budgets for sites in the Impact Study were statistically significantly higher, ranging from $100,000 to $1,000,000, with an average of $277,000, than for sites from the representative sample, where budgets ranged from $65,000 to $1,738,000, with an average budget of $232,000.

**Exhibit 3.1**

**Grantee Characteristics**

| Measured Characteristic | Impact Study Sites | | Representative Sample of Grantees | | p-value to Test Difference |
|---|---|---|---|---|---|
| | Mean/Percentage | Standard Deviation[a] | Mean/Percentage | Standard Deviation | |
| *Years operating (mean)* | 6.1 | 5.6 | 4.9 | 5.5 | .001* |
| *Auspice (%)* | | | | | |
| Nonprofit/Community-based organization or Faith-based organization | 65.6% | | 71.0% | | .013* |
| School or school district | 34.4% | | 29.0% | | .013* |
| *Program size (mean)* | | | | | |
| Annual school-based mentoring program budget | $276,800 | $208,326 | $232,351 | $233,961 | .001* |
| Number students served | 217 | 199.5 | 146 | 289.3 | .000* |

*two-tailed significance, p<.05

Impact Study Sites: N = 32; Missing data ≤ 3%

Representative Sample of Grantees: N = 100; Missing data ≤ 9%

[a] Standard Deviations are only reported for Means or Mean Percents.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Grantee Survey, Spring 2006 and Spring 2007.

Impact Study sites were also statistically significantly larger in terms of the numbers of students they served, with an average of 217 students (per site) compared to 146 students for sites from the representative sample. These differences are more notable when we consider the number of sites operating small programs (fewer than 50 students) and large programs (250 or more students). In addition to comparing the average number of students served by the Impact Study sites with the average number of students served at sites from the representative sample of grantees, we also assessed differences in program size by considering differences between the numbers of grantees that operate programs of specific sizes. Two-fifths (40 percent) of sites in the representative sample served fewer than 50 students, compared to 9 percent of the Impact Study sites, and this difference was statistically significant. We observed similar differences between the Impact Study and representative sample sites on the other end of the size continuum: approximately a third (34 percent) of the Impact Study sites served 250 or more students compared to 13 percent of the sites in the representative sample, and, again, this difference was statistically significant.

These results demonstrate that grantees in the Impact Study sample tended to have more experience operating mentoring programs, to serve larger numbers of students, to have larger budgets, and also were more likely to be run by school districts (and less likely to be nonprofit/community-based organizations or faith-based organizations) than the typical grantee. Organizations selected to participate in the Impact Study may have been more established and ready to begin operations early in the school year than other grantees because sites in the study had to commit to recruit and survey at least 60 students by November of the study year and to match one-half of those students with mentors in the program either by that time or shortly thereafter. Small grantees may have been less likely than larger ones to participate because they may have had difficulty meeting these requirements.

Including grantees that were likely to be up and running relatively early in the school year also may explain why greater percentages of schools and school districts were included in the Impact Study sample than in the representative sample of grantees. Because they may have been better situated to identify and recruit students than nonprofit/community-based organizations or faith-based organizations, they may also have been better able to have programs functional in time to participate in the baseline data collection that was a part of the Impact Study and to recruit at least 60 students by November of the respective school year. Being a school or a school district also may have rendered them better able to facilitate various components of the evaluation, such as group survey administration at schools and student record data abstraction, than outside community organizations.

**Reported Program Focus**

Exhibit 3.2 summarizes grantees' reports of the issues on which they are extremely focused. In line with the absolute priority of meeting the academic and social needs of students, the overwhelming majority of grantees reported being extremely focused on a number of factors associated with students' positive development, including: improving students' academic performance (91 percent of the Impact Study sites versus 95 percent of the representative sample, a statistically significant difference), and building students' self-esteem, which was a focus for 84 percent of grantees from both the Impact Study and the representative sample; providing unspecified general guidance, which was a focus for nearly three-quarters in both groups of all sites (72 and 73 percent, respectively, for Impact Study and representative sample grantees), and relationship building, which was a focus for 62 and 63 percent, respectively, for Impact Study and representative sample sites.

In contrast, relatively few Impact Study sites reported being extremely focused on risk avoidance, although sites in the representative sample were more heavily focused on this issue. Thirty-eight percent of the Impact Study sites reported being extremely focused on risk avoidance compared to 60 percent of the sites from the representative sample. This statistically significant difference may be explained by the fact that a statistically significantly larger percentage of sites in the representative sample were nonprofit/community-based organizations or faith-based organizations (and not schools or school districts) and therefore may have been more likely to be focused on risk and risk avoidance (as opposed to academic outcomes). Both groups of grantees were also less intent on increasing community engagement. Less than a third of Impact Study sites (28 percent) and sites from the representative sample (32 percent) reported that this was a major focus.

**Exhibit 3.2**

**Grantee-Reported Program Focus**

| Measured Characteristic | Impact Study Sites (Percentage) | Representative Sample of Grantees (Percentage) | p-value to Test Difference |
|---|---|---|---|
| *Extremely focused on:* | | | |
| Academics [a] | 90.6 | 95.0 | .002* |
| Self-esteem | 84.4 | 84.0 | .357 |
| Providing unspecified, general guidance | 71.9 | 73.0 | .468 |
| Relationship building (with individuals other than mentors) [b] | 62.5 | 63.0 | .512 |
| Risk avoidance [c] | 37.5 | 60.0 | .000* |
| Increasing community engagement | 28.1 | 32.0 | .173 |

*two-tailed significance, p<.05

Impact Study Sites: N = 32; No Missing data

Representative Sample of Grantees: N = 100; No Missing data

[a]  Sites were considered extremely focused on academics if they selected "extremely focused" for any one of the following four items: improving mentees' attitudes towards school, improving mentees' academic performance in school, improving mentees' attendance, and improving the likelihood that mentees will not drop out of school before graduating from high school.

[b]  Sites were considered extremely focused on relationship building (with individuals other than mentors) if they selected "extremely focused" for any one of the following three items: improving mentees' relationships with their parents/other caregivers, improving mentees' relationships with other adults in authority (teachers, principals, probation officers, etc.), and improving mentees' relationships with peers.

[c]  Sites were considered extremely focused on risk behaviors if they selected "extremely focused" for any of the following five items: increasing mentees' ability to refrain from getting involved in gangs, increasing mentees' ability to refrain from engaging in violent activities, increasing mentees' ability to refrain from engaging in criminal activities, increasing mentees' ability to refrain from using drugs/alcohol, increasing mentees' ability to refrain from high-risk sexual behaviors.

*Source:*   Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Grantee Survey, Spring 2006 and Spring 2007.

## Students Served by the Grantees

### *Student Demographics*
Both grantees in the Impact Study sample and grantees representing the universe of Student Mentoring Program sites reported serving students from a variety of ethnic and racial backgrounds. Exhibit 3.3 indicates the ethnic and racial background of students in the Student Mentoring Program. It also indicates students' grade levels and gender.[43]

---

[43]  To compare the ethnicities of students served in the Impact Study sites and other grantees not in the Impact Study, we relied on grantee reports of the percentages of the ethnicities of students served. Although actual data on student ethnicity from student records would probably have been more accurate, because we do not have comparable data for students not in the Impact Study sites, we relied on grantee reports of student ethnicity so that we could compare the two groups of sites using comparable data.

**Exhibit 3.3**

**Grantee Characteristics: Students Served (Unweighted Mean Percentages)**

| Measured Characteristic | Impact Study Sites | | Representative Sample of Grantees | | p-value of Differences |
|---|---|---|---|---|---|
| | Mean Percentage | Standard Deviation[a] | Mean Percentage | Standard Deviation | |
| *Student race/ethnicity* [b] | | | | | |
| American Indian or Alaskan Native | 3.8 | 14.4 | 4.2 | 15.4 | .613 |
| Asian | 2.3 | 5.8 | 0.9 | 2.2 | .000* |
| Black or African American | 41.2 | 34.5 | 36.7 | 36.0 | .123 |
| Hispanic or Latino | 29.1 | 28.5 | 22.5 | 29.5 | .002* |
| Native Hawaiian or other Pacific Islander | 0.4 | 1.1 | 0.1 | 0.7 | .000* |
| White | 23.1 | 19.6 | 34.4 | 33.6 | .000* |
| *Student gender* | | | | | |
| Female | 57.3 | 15.8 | 51.8 | 14.2 | .000* |
| *Student grade* [c] | | | | | |
| Grades 4-5 | 41.7 | 24.7 | 38.5 | 28.9 | .057 |
| Grades 6-8 | 44.4 | 24.3 | 46.0 | 30.6 | .279 |
| Other grades | 13.9 | 24.7 | 15.5 | 19.0 | .362 |

*two-tailed significance, p<.05

Impact Study Sites: N = 32; Missing data = 0%

Representative Sample of Grantees: N = 100; Missing data ≤ 4%

[a] Standard Deviations are only reported for Means or Mean Percents.

[b] 25% of the Impact Study sites and 20% of the representative sample of grantees reported not systematically collecting this information.

[c] 28% of the Impact Study sites and 14% of the representative sample of grantees reported not systematically collecting this information.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Grantee Survey, Spring 2006 and Spring 2007.

Exhibit 3.3 demonstrates that grantees in the Impact Study and representative sample served similar percentages of students who were American Indians and Alaskan Natives (4 percent for both groups), blacks or and African Americans (41 percent for the Impact Study sites and 37 percent for sites in the representative sample). Impact Study sites tended to serve higher percentages of Latino students (29 versus 23 percent), Asian and Native Hawaiian or other Pacific Islander students (3 versus 1 percent) and lower percentages of white students (23 versus 34 percent) than grantees in the representative sample. All of these differences were statistically significant. Finally, grantees in the Impact Study served a statistically significantly higher percentage of girls compared to grantees in the representative sample (57 versus 52 percent, respectively).

Although most students served were in the age range targeted by the authorizing legislation (i.e., students in grades 4–8), on average, 14 percent of students served by the Impact Study grantees were not in this age range compared to 16 percent of students in the representative sample of grantees (this difference was not statistically significant). Students not in the targeted age range were not eligible to participate in the Impact Study.

### *Student Risk Factors*

The program legislation called for grantees to serve children with the greatest need as defined by "at risk of educational failure, dropping out of school, or involvement in criminal or delinquent activities, or who lack strong positive role models."[44] As Exhibit 3.4 illustrates, grantees in both the Impact Study and representative samples, following the directives set forth by the legislation, cited the same three risk factors as most common for students they served. Three-quarters of Impact Study grantees and 70 percent of grantees in the representative sample identified "students had few or no positive adult role models" as one of the top three risk factors for students in their programs. Having self-esteem problems was another common risk factor for students in both groups (56 percent of Impact Study sites and 62 percent of sites in the representative sample). While both groups of grantees also indicated that failing in school was among the top three risk factors, there was a statistically significant difference in terms of how often grantees cited this problem. Fifty-three percent of the Impact Study sites compared to 43 percent of grantees in the representative sample identified this risk factor as one of the top three for students in their programs.

**Exhibit 3.4**

**Grantee Characteristics: Top Three Perceived Student Risk Factors**

| Measured Characteristic | Impact Study Sites (Percentage) | Representative Sample of Grantees (Percentage) | p-value of Differences |
|---|---|---|---|
| *Perceived risk factors* | | | |
| Student has few/no positive adult role models. | 75.0 | 70.0 | .082 |
| Student has self-esteem problems. | 56.3 | 62.0 | .101 |
| Student is failing in school. | 53.1 | 43.0 | .000* |

*two-tailed significance, p<.05

Impact Study Sites: N = 32; No Missing data

Representative Sample of Grantees: N = 100; No Missing data

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Grantee Survey, Spring 2006 and Spring 2007.

### Summary of Differences in Grantee Characteristics

In summary, although grantees in the Impact Study were similar to the grantees in the representative sample in many respects, there were several statistically significant differences between the two sets

---

[44]   No Child Left Behind Act, Section 4130 – Mentoring Programs, 2002.

of grantees.[45] The statistically significant differences were the kind of organization running the mentoring program, the number of students mentored, program budget, grantee experience, and program focus on risk avoidance. These results not only confirm observable differences in important grantee characteristics, but they further suggest that impact findings based on the Impact Study grantees may be limited in their generalizability (i.e., results are not necessarily generalizable to the overall population of Student Mentoring Program grantees).

## 3.3   Characteristics of the Mentors of Impact Study Students

This section discusses the characteristics of the individuals who mentored students assigned to the treatment group in the Impact Study.

**Mentor Demographics**

As indicated in Exhibit 3.5, 72 percent of the mentors who participated in the Student Mentoring Program were women. Mentors ranged in age from 12 to 82, with an average age of approximately 32. This relatively young average age reflects the fact that 41 percent of the mentors in our study were age 22 and under. Mentors came from a variety of racial and ethnic backgrounds with approximately two thirds (66 percent) of mentors being white.  Twenty-nine percent of mentors were black or African American. Less than 10 percent of the mentors were American Indian or Alaskan Native, Asian or Native Hawaiian or other Pacific Islander.[46] One out of every ten mentors reported their being of Hispanic or Latino ethnicity. Almost all mentors (95 percent) were native English speakers. Approximately one-third (34 percent) of the mentors were married or living with a partner, while 36 percent reported having children.

**Mentor Education and Employment**

As noted in the previous section, 41 percent of mentors were ages 22 and under. Based on a set of questions from the Mentor Survey on their educational status and attainment, we calculated that 31 percent of mentors were full-time students who had completed high-school.  Another 7 percent of mentors were full-time students who had *not* completed high-school.[47]

These findings parallel mentor reports of the highest level of education they had completed at the time they were surveyed (Exhibit 3.6). Ten percent had completed some high school; 8 percent reported a high-school degree or GED as their highest level of education; and 82 percent reported having some college or post-secondary training or more. Eighty-eight percent of mentors were either full time students or employed full- or part-time. Thirty-nine percent of mentors were full-time

---

[45]   These differences could be attributable to the requirements imposed on grantees for participation in the Impact Study.

[46]   Note that these totals equal more than 100 percent because mentors could select more than one race category (e.g., both white, and black or African American).

[47]   The Mentor Survey did not specifically ask mentors if they were high-school or college students, therefore, we inferred that mentors who indicated that they were full-time students and had not completed high school were high-school students, and that mentors who were full-time students who had completed high school were either college, technical school or graduate school students.

students and approximately half (49 percent) were employed full- or part-time. An additional 12 percent were either retired or not employed outside of the home.

**Exhibit 3.5**

**Mentor Demographics**

| Measured Characteristic | Percentage |
|---|---|
| *Demographic information* | |
| Gender: Percent female | 71.8% |
| Age (in years) | 32.1[a] |
| | (10.2) |
|     18 or younger | 17.5% |
|     19-22 | 23.4% |
|     23-64 | 56.1% |
|     65+ | 3.0% |
| Ethnicity: Percent Hispanic/Latino | 10.0% |
| Race [b] | |
|     White | 65.8% |
|     Black or African American | 29.2% |
|     Asian | 4.8% |
|     American Indian or Alaskan Native | 3.3% |
|     Native Hawaiian or other Pacific Islander | 1.0% |
| Native language: Percent English speakers | 94.6% |
| *Family information* | |
| Mentors married or living with partner | 33.8% |
| Have children | 35.5% |
| N = 974 | |
| Missing data ≤ 6% | |

[a]  Mean number of years (Standard deviation).

[b]  Percentages add up to more than 100 because mentors could select multiple races.

*Source:*  Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – Mentor Survey, Spring 2006 and Spring 2007.

**Exhibit 3.6**

**Mentor Education and Employment**

| Measured Characteristic | Percentage |
|---|---|
| *Student status* [a] | |
| High-school students:  Full-time students, who *have not* completed high school | 7.0% |
| College students:  Full-time students, who *have* completed high school | 31.3% |
| *Highest educational attainment* | |
| Some high school | 10.3% |
| HS degree or GED | 7.6% |
| Some college or post-secondary school training or more | 82.1% |
| *Employment status* | |
| Employed, full- or part-time | 48.9% |
| Full-time students | 39.1% |
| Other (Retired or not employed outside of home) | 11.9% |
| N = 974 | |
| Missing data ≤ 2% | |

[a]  High school/GED completion is unknown for some mentors who were full-time students. As a result, percentages total less than 39.1 percent (the percentage reported for full-time students under employment status in this table).

*Source:*   Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – Mentor Survey, Spring 2006 and Spring 2007.

### Mentor Prior Experience

As indicated in Exhibit 3.7, 76 percent of mentors reported having had "some" or "a lot" of  contact with students in grades 4–8 prior to their participation in the Student Mentoring Program,  Despite that fact, a smaller percentage had prior experience mentoring. Less than one-third (31 percent) of mentors reported that they had volunteered as a mentor prior to their participation in the program.

**Exhibit 3.7**

**Mentors' Prior Experience**

| Measured Characteristic | Percentage |
|---|---|
| *Self-reported experience mentoring* | 30.9% |
| *Self-reported contact with students in grades 4-8* | |
| None | 6.7% |
| Very little | 17.4% |
| Some | 37.2% |
| A lot | 38.7% |

N = 974

Missing data ≤ 1%

*Source:*  Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – Mentor Survey, Spring 2006 and Spring 2007.


## 3.4   Characteristics of Program Delivery

**Mentor/Student Relationships**

*Race/Ethnicity and Gender Matches*

Because some research suggests that mentoring may be particularly effective when mentors and students are from the same or similar racial backgrounds,[48] we determined the percentages of matches in the Student Mentoring Program in which the mentor and student were from the same ethnic/racial groups. Fifty-five percent of the matches in our study were between individuals who had the same racial status; of the cross-race matches (45 percent of all matches), 82 percent were between a white mentor and minority student, 12 percent were between mentors and students from different minority groups, with the remaining 5 percent between white students and minority mentors.

We also explored the extent to which mentors and students were of the same gender.[49] Approximately four-fifths (81 percent) of matches in our study were between students and mentors of the same gender. Of the 19 percent of matches that were cross-gender, 92 percent involved male students with female mentors. Exhibit 3.8 summarizes our findings about gender and race in mentoring relationships.

*Relationship Length*

Several studies suggest that characteristics and quality of the mentoring relationship are the most consistent predictors of impact (Herrera, et al., 2000; Rhodes, et al., 2000). Specifically, there is some evidence (Grossman, et al., 1999; Herrera, et al., 2000; DuBois and Neville, 1997; Nakkula and Harris, 2005) that characteristics of the relationship such as duration of the mentoring relationship

---

[48]   See Sanchez and Colon (2005), 191-204.

[49]   The literature is mixed about the relative efficacy of cross versus same gender matches. Most programs do, however, attempt to match students and mentors of the same gender. For more discussion of the role of gender in mentoring, see Bogan and Liang (2005).

affect the level and type of impacts on student participants. For example, in the Big Brothers/Big Sisters experimental evaluation of community-based mentoring (Tierney and Grossman, 2000), students whose matches lasted 12 months or more skipped fewer classes and had higher grades and less drug initiation at follow-up than those whose matches lasted 3–6 months. In addition, correlational analyses from earlier studies have found associations between negative student outcomes and early termination of matches or inconsistent attendance on the part of mentors (Slicker and Palmer, 1993; Karcher, 2005; Portwood and Ayers, 2005). Finally, although these results were non-experimental, Herrera (2004) found that students in matches lasting nine months or more showed statistically significant improvements in several of the areas targeted in the Student Mentoring Program legislation, including relationships with peers, classroom behavior, numbers of referrals to principals' offices and fighting.

**Exhibit 3.8**

**Race/Ethnicity and Gender Matching in Mentoring Relationships**

| Measured Characteristic | Percentage of All Mentors | Percentage of Subgroup of Mentors |
|---|---|---|
| *Minority Status*[a, b] | | |
| Mentor and student were same race/ethnicity | 54.7% | |
|   White mentor/white student | | 39.7% |
|   Minority mentor/same minority student | | 60.3% |
| Mentor and student were different race/ethnicity | 45.3% | |
|   White mentor/minority student | | 82.3% |
|   Minority mentor/different minority student [c] | | 12.4% |
|   Minority mentor/white student | | 5.3% |
| *Gender* | | |
|   Mentor and student were same gender | 80.8% | |

N = 1,050

Missing data ≤ 12%

[a] In cases where schools did not provide data on student race for individual students, we checked the website www.greatschools.net to see if schools had 95 percent or more of students from the same racial background. When that was the case, we imputed that particular race for students from those schools.

[b] Percents are approximate because of two differences in the format of supplied data. First, in student records, schools reported only one variable on race/ethnicity, but in the Mentor Survey, race and ethnicity were separate items. For this analysis, if a student indicated that he was white while his mentor reported that he was white and Hispanic/Latino, we considered them to be the same ethnicity, and counted the match as a white/white match. Furthermore, the Mentor Survey has a separate category for Native Hawaiian or Pacific Islander, while school district data does not. Since schools are most likely to categorize these students' race as Asian, we considered student and mentor a race-match if a mentor identified as a Native Hawaiian or Pacific Islander and student was identified as Asian.

[c] Pairs in this category could include, for example, a black or African American mentor with an Asian student.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – Mentor Survey, Spring 2006 and Spring 2007.

We therefore collected information on match duration for students and mentors in the Impact Study. Exhibit 3.9 provides an overview of the student/mentor relationship length for the entire treatment group sample, including students who were never matched with mentors; students whose mentors never completed the Mentor Survey, students who were matched but never met with their mentors, as well as students who met with mentors as intended. Exhibit 3.9 indicates that 17 percent of the treatment group did not receive any mentoring from the program. This includes 14 percent of students in the treatment group who were never matched with mentors and another 3 percent who were matched with mentors, but never actually met. Forty-four percent of the matches made through the Student Mentoring Program lasted for at least six months and another one-fifth (20 percent) lasted between three and six months. Another 6 percent of students were matched and met with mentors for fewer than three months. For 13 percent of the sample, information on relationship length was not available due to non-response.

**Exhibit 3.9**

**Extent of Interactions between Students and Mentors for Treatment Group Sample**

| Measured Outcome | Percentage |
| --- | --- |
| Students never matched | 14.0% |
| Students matched but mentor reported they never met | 3.0% |
| Students met with mentor(s) for [a]: | |
| Less than 3 months | 5.8% |
| 3 – 6 months | 20.3% |
| More than 6 months | 43.6% |
| Students matched but mentor never responded (to entire survey or to question of whether they met) | 13.1% |

N = 1,272

No missing data

[a]  Match lengths are aggregated for students who met with more than one mentor.

*Source:*   Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – Mentor Survey, Spring 2006 and Spring 2007.

*Frequency and Duration of Student/Mentor Meetings*

As shown in Exhibit 3.10, of those mentors completing the survey and who met with their students (95 percent), most (87 percent) met one-on-one with students as opposed to the remaining mentors who mentored two or more students at a time, by themselves or with one or more other mentors. Mentors who met with their students did so an average of 4.4 times per month and meetings lasted 1.1 hours (on average). The average mentoring relationship for mentors who were matched and met with their students was 5.8 months.[50]  Programs took, on average, a total of 81 days to match their students

---

[50]   These data reflect month-based estimates based on mentor reports of the beginning and end months of mentoring. They are also limited by the fact that the Mentor Survey was administered approximately one month before the end of the program year. Our survey thus asked mentors to indicate the month they stopped mentoring *or* expected to stop mentoring their students, and they could not report any match end dates past June of the school year.

from the beginning of the school year (set to September 1). From the time of random assignment, the average length of time to make a match was 35 days.[51]

---

**Exhibit 3.10**

**Student: Mentoring Sessions and Amount of Mentoring Received**

| Measured Characteristic | Mean or Percentage | Standard Deviation[a] |
|---|---|---|
| Met one-on-one with student(s) [b] | 87.4% | |
| Meeting frequency: Number of in-person contacts per month [c] | 4.4 | 2.8 |
| Meeting duration: Average length of meeting [c] | 1.1 hours | 0.7 |
| Duration of mentoring relationship [c] | 5.8 months | 2.3 |
| Average amount of time to match students from beginning of school year [d] | 80.7 days | 38.7 |
| Average amount of time to match students from time of random assignment [d] | 36.7 days | 35.1 |
| Missing data ≤ 9% | | |

[a]   Standard Deviations are only reported for Means or Mean Percents.

[b]   This item is based on the mentors who answered our survey and also met with their students. N for this item is 974. Source for this item is the Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – Mentor Survey General Questions, Spring 2006 and Spring 2007.

[c]   This item is based on the mentors who answered our survey and also met with their students. N for this item is 1,050. Source for this item is the Mentor Survey Student Specific Questions, Spring 2006 and Spring 2007.

[d]   N for this item is 1,131. Source for this item is the Student Tracking File.

*Source:*   Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – Mentor Survey, Spring 2006 and Spring 2007, Student Tracking File.

---

*Student/Mentor Activities*

As illustrated in Exhibit 3.11, mentors and students spent time engaged in a variety of activities. In terms of activities potentially linked to student outcomes, the most commonly reported activity for mentors and students who met was discussing students' relationships with others, including parents, peers, teachers and other adults in authority. This activity is targeted in the Student Mentoring Program legislation, which indicates improved interpersonal relationships as one of its goals. Fifty-two percent of mentors reported engaging in this activity "most of the time" or "almost always."

The legislation also targets improved academic performance and participation. Forty-three percent of mentors who met with students reported working on academics with their students either "most of the time" or "almost always," and 48 percent reported discussing their students' plans for the future (including completing high school) either "most of the time" or "almost always." In addition, 23 percent reported having discussions about risk behaviors with their students on a frequent basis.

---

[51]   These data were calculated from the study's student tracking file, which included the date of random assignment for every student and date of match for every treatment group student (and for "crossover" control group students).

**Exhibit 3.11**

**Student/Mentor Activities**

| Activities that Were Part of Meeting: [a] | Never | Sometimes | Most of the time or Almost Always |
|---|---|---|---|
| Worked on academics [b] | 21.3% | 36.2% | 42.5% |
| Discussed student's plans for future, including completing high school | 8.0% | 44.0% | 47.9% |
| Discussed student's relationships with peers, parents, teachers, or other authority figures | 3.4% | 44.5% | 52.1% |
| Discussed risk-behaviors [c] | 39.0% | 38.3% | 22.8% |
| Engaged in community service with youth | 77.4% | 18.7% | 3.9% |

N = 1,050

Missing data ≤ 8%

[a]  Based on mentor reports.

[b]  Respondents are counted if they answered either: Worked on academic skills or Worked on homework.

[c]  Respondents are counted if they reported discussing any of the following risk behaviors: Alcohol/drug use, Engaging in violence/criminal activities, Gang involvement, or High-risk sexual activity.

*Source:*   Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – Mentor Survey, Spring 2006 and Spring 2007.


*Mentor and Student Perceptions of the Relationship*

The mentor/student relationship is intended to lead to an array of positive student behaviors supporting intended long-term program impacts. Not all mentoring relationships that begin, however, will last or be sufficiently positive to generate intended outcomes. To assess attributes of relationships in the Student Mentoring Program, we considered both student and mentor perceptions of the relationships.

As shown in Exhibit 3.12, 78 percent of mentors who met with students in the Student Mentoring Program reported that their relationships with students were somewhat or extremely positive. The overwhelming majority (94 percent) reported enjoying the time with their students on a frequent basis (i.e., either "most of the time" or "almost always."), while 49 percent of mentors reported that their student(s) frequently confided in them  In contrast, 8 percent of mentors reported that it was frequently hard for them to engage their students in conversation.

**Exhibit 3.12**

**Mentor/Student Relationships: Mentor Perspective**

| Measured Characteristic | Percentage of Mentors Who Met with Students |
|---|---|
| Mentors perceived their relationships with students as somewhat or extremely positive [a] | 78.2% |
| Mentor/Student Relationship Scale (mentor self-report) items [b] | |
| You enjoyed the time with your student | 93.7% |
| Your student confided in you | 49.0% |
| It was hard for you to engage your student in conversation | 8.0% |

N = 1,050

Missing data ≤ 7%

[a]  N for this item is 974.

[b]  Percent based on responses of Most of the time/Almost always to each item.

*Source:*   Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – Mentor Survey, Spring 2006 and Spring 2007.

As shown in Exhibit 3.13, students also had positive perceptions of their mentors.  Most students who met with mentors answered that they felt that they could trust their mentors (89 percent), could rely on their mentors to listen to them (86 percent) and help solve their problems (90 percent).

**Exhibit 3.13**

**Mentor/Student Relationships: Student Perspective**

| Measured Characteristic | Percentage of Students Who Met with Mentors |
|---|---|
| Mentor/Student Relationship Scale (student self-report) items [a] | |
| I feel that I can trust my mentor | 88.6% |
| When something is bugging me, my mentor listens to me | 86.4% |
| My mentor has good ideas about how to solve problems | 90.0% |

N = 1,005

Missing data ≤ 9%

[a,]  Percentages for these items are based on responses of Sort of True or Very True.

*Source:*   Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – Student Survey, Spring 2006 and Spring 2007.

**Mentor Support and Supervision**

Student Mentoring Program grantees were required to provide mentors with a variety of supports, including pre-match screening and orientation and ongoing support for matches. According to mentor reports, most grantees complied with these requirements. As illustrated in Exhibit 3.14, 89 percent of mentors reported having either a background or reference check conducted pre-match, despite the fact that such checks are required as a condition of the grant. Because only mentors and not grantees were asked this question, it is possible that some mentors were simply unaware (or had forgotten) that a background or reference check was conducted by the grantee.[52] In addition, 96 percent of mentors reported receiving some form of pre-match training or orientation. Those mentors that received pre-match training or orientation received an average of 3.4 hours of this support. Forty-one percent of mentors reported that ongoing training was available after they had begun meeting regularly with their students.

**Exhibit 3.14**

**Mentor Support and Supervision**

| Measured Characteristic | Percentage of all Mentors | Mean or Percentage of Subgroup of Mentors (s.d.) |
|---|---|---|
| *Pre-match screening, training, orientation* | | |
| Background or reference check conducted | 89.4% | |
| Received some sort of pre-match training/orientation | 96.3% | |
| Number of hours of pre-match training/orientation, of mentors who received any | | 3.4 hours (4.1) |
| *Ongoing mentor support* | | |
| Ongoing training available | 40.8% | |
| Ongoing supports provided | 93.6% | |
| *Type of resources/supervision provided* | | |
| Supervised mentor/student meetings | 51.0% | |
| Access to social workers or program staff | 62.3% | |
| Access to listservs or other online forums, in-person get-togethers with other mentors, or other support | 57.8% | |
| Mentors talked with program supervisor about how things were going | 92.8% | |
| Reporting to program staff was required or strongly encouraged | | 72.0% |

N = 974

Missing data ≤ 5%

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – Mentor Survey, Spring 2006 and Spring 2007.

---

[52] In fact, all 32 grantees in the Impact Study indicated that they required some form of background screening before matching mentors with students.

The majority of mentors (94 percent) reported having access to some kind of ongoing supports, but the extent and content of those supports varied. Fifty-one percent of mentors reported that their meetings were supervised by program staff, and 62 percent reported having access to social workers or program staff (even if meetings were not supervised). Of those mentors (93 percent) who met or talked with mentoring supervisory staff about their matches, 72 percent indicated that these check-ins were either required or strongly encouraged. In addition, 58 percent of mentors reported the availability of a variety of additional supports, including on-line discussion forums and listservs, informal get-togethers for participating mentors, and other supports.

## 3.5   The Treatment Contrast

Finally, in Exhibit 3.15, we report on students' report of receipt of mentoring from *all* sources, not just ED Student Mentoring Programs. Since we obtained this information for both treatment and control groups, we are able to make comparisons about the relative receipt of mentoring for both groups. It needs to be emphasized here, that the recruiting process of both sites and students into the study made clear that students assigned to the control group, although ineligible to receive mentoring through the Student Mentoring Program, were not excluded from seeking other mentoring or similar services in the community.[53]

In the spring follow-up interviews, 86 percent of treatment group students reported receiving mentoring through any program during the past school year compared to 35 percent of the control group students who accessed ED or other mentoring services; this difference was statistically significant. Treatment group students also met more frequently with their mentors than control group students. For example, approximately a third (34 percent) of the control group students met with their mentors once a month or less compared to 15 percent of the treatment group students, a statistically significant difference. In contrast, 85 percent of students in the treatment group receiving mentoring met with their mentors at least two times per month compared to approximately two-thirds (66 percent) of the control group students who reported being mentored (a statistically significant difference).

---

[53]    However, according to information obtained from each program, we determined that a total of 3 percent of all students originally randomly assigned to the control group did eventually get mentoring through the ED Student Mentoring Program.

**Exhibit 3.15**

**Contrast in Receipt of Mentoring[a]**

| Measured Outcome | Treatment Group Percentage | Control Group Percentage |
|---|---|---|
| Student-reported receipt of mentoring through ***any*** program | 85.7% | 35.0% [b] |
| Student-reported frequency of mentor/student meetings across all programs, of students who received mentoring through any program [c] | | |
| Once a month or less | 14.0% | 34.0% |
| 2 times a month or more | 86.0% | 66.0% |

N = 1,173 for treatment, 1,199 for control

Missing data ≤ 15%

[a]  All differences between Treatment and Control Groups in this exhibit are statistically significant, p<.05.

[b]  This figure includes 3 percent of those students assigned to the control group who crossed over into the treatment group and received mentoring from a study grantee during the study follow-up period.

[c]  N for this item is 950 for treatment, 359 for control.

*Source:*  Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – Student Survey, Spring 2006 and Spring 2007.

## 3.6    Summary of Descriptive and Program Delivery Findings

The Impact Study grantees, in general, appeared to be providing mentoring consistent with what was outlined in the legislation and what was recommended by the program office.  Grantees reported, for the most part, serving the appropriate students in terms of age and risk factors as specified in the legislation, and targeting both the academic and social needs of students as prioritized by OSFDS.  Mentors also reported having available to them a variety of supports from their programs as required by program guidelines, ranging from pre-match training to supervised meetings.   In addition, both mentors and students reported having positive feelings about their mentoring relationships. However, there were a number of issues with program delivery that deserve mention:

- Approximately one out of every ten mentors reported that they never underwent a background or security clearance even though grantees are required to conduct these clearances on all mentors as a condition of their grant.  All grantees, however, reported that they conducted these checks prior to matching a mentor with a student.

- Seventeen percent of the students randomly assigned to the treatment group never received mentoring from the program.  The majority (i.e., 82 percent) of these students were never matched with a mentor.  This occurred for a number of reasons, including the site's inability to find an appropriate match for the student, or refusal on the part of the student or on the part of the student's parent or guardian to consent to student mentoring.  Although the percentage of unmatched students in this study is within the range of past

experience engaging mentors in randomized impact studies,[54] the inability to match students may have led to some dilution of the anticipated treatment effect.  In addition, another possible dilution factor is the finding that of those students who met with their mentors, 13 percent did not receive mentoring on a 1:1 basis.[55]

- Eighteen percent of mentors were 18 or younger and an additional 23 percent were college-age; thus, they may not have had appreciably more life experience than the students they were mentoring.   In contrast, Herrera et al. (2007) reported that half of the mentors in the BBBS study were 18 years old or younger, with an additional 17 percent 19 to 24 years old.  She also notes that while using high school students increases the number of students a program can serve (and hence, increase the probability that a given student will be matched), recruiting mentors from this age group necessarily limits the length of matches given students' inability to commit beyond a semester or school year.

- Although the majority (i.e., 76 percent) of mentors reported having had "some" or "a lot of" contact with children in the target age group (i.e., grades 4-8), 31 percent reported previous experience mentoring.

- The programs took a relatively long time to match students and have them meet their mentors.  The average time between the start of the school year and the date students were matched with a mentor) was 81 days (range: 38 to 132 days at the site level).  On average, there was a lag of 37 days between the date of random assignment and the time when the student was matched (ranging at the site level from 0 to 97 days).  The date of the first meeting usually occurred at or shortly after the date the match was made. This slow startup has been previously cited as common in school-based mentoring. Karcher (2008) noted a 2- to 3-month lag in getting students matched from the beginning of the school year, consistent with earlier findings from Hansen (2005) and Herrera et al. (2000).

- The average relationship was 5.8 months in length, which is less than the recommended 12 months of time by Rhodes (2002) and others in the mentoring field,[56] and was in part a consequence of the constraints of the school calendar, limiting the maximum length of a match to approximately 9 months, as well as the previously cited length of time required to make a match. This finding is consistent with previous research (cf. Herrera et al., 2007, Karcher, 2008) and correlational findings from some studies (Rhodes, 2002, Rhodes, Reddy, Roffman and Grossman, 2005), have suggested associations between early termination or short-lived matches and negative outcomes for students.

- Ninety percent of grantees reporting being "extremely focused" on academics, while 43 percent of the mentors reported working frequently on academics, with an additional 21 percent reporting **never** working on academics.  In terms of social needs, 40 percent of

---

[54]  For example, Herrera et al. (2007) reported 7 percent of the students unmatched in the first year of the BBBS study and 5 percent by the second follow-up. In a 2000 study on community-based mentoring, Tierney and Grossman found that 22 percent of the youth randomly assigned to the treatment group were unmatched.

[55]  Although the fact that not all students received 1:1 mentoring may potentially limit program impact, we are unable to assess within the experimental design whether results may differ for those students receiving 1:1 mentoring given our inability to indentify their respective counterfactuals in the control group.

[56]  This recommendation, however, has typically been applied to community-based mentoring programs.

grantees reported being extremely focused on risk-avoidance.  From the mentor perspective, 52 percent reported frequently discussing relationships with students.  However, 39 percent of mentors reported ***never*** discussing risky behaviors.[57]

These findings, taken as a whole, show that program grantees participating in the evaluation generally adhere to the intent of the legislation and direction from the OSDFS, which mandates targeting a number of outcomes without much specificity regarding how to deliver the mentoring.   However, beyond what was specified in the legislation and OSDFS, there were a number of issues with program delivery.  For example, 17 percent of students assigned for mentoring never actually met with their mentors and the actual amount of mentoring provided by program grantees participating in the evaluation was at a fairly low level of intensity in terms of the total number of hours that students actually had contact with their mentors, due to the limited duration of the mentoring relationship, 5.8 months on average.  Mentors also tended to be young (42 percent were age 22 and under) and had little previous experience with mentoring (31 percent).  These issues with program delivery are consistent with what has been found in previous research on school-based mentoring programs.

---

[57]   This last finding could possibly be attributed in part to the low incidence of high-risk behaviors for this age group. Recall the previously cited finding in this chapter that one-fourth (25) percent of students presented any indication of high-risk behaviors at baseline.  Still, they may have been contemplating them and, in any case, these issues become more important as students age during the mentoring year.

# Chapter 4: Impacts on Students

This study was designed to determine the impact of the Student Mentoring Program on students in 32 purposively-selected programs funded by ED in 2004 and 2005. In this chapter we describe the results of the analysis of the impact of programs on students in areas that this type of intervention hopes to influence: increasing school engagement and improving academic achievement, lowering high-risk and delinquent behavior, and improving interpersonal relationships with adults, personal responsibility, and community involvement. The chapter concludes with an overview of all findings in the report.

## 4.1    Overview of Impact and Correlational Analyses

We found that the Student Mentoring Program had no overall statistically significant impacts on the key student-level outcomes measured in this study after adjusting for multiple comparisons. However, there were some statistically significant differences in impacts across subgroups. More specifically:

- Overall Impact Findings:
    - Students in the treatment group did not report statistically significant differences in interpersonal relationships, personal responsibility, and community involvement at the end of the spring school term relative to students in the control group.
    - After controlling for multiple comparisons, students in the treatment group did not exhibit statistically significant differences in academic achievement or school engagement relative to students in the control group.
    - After controlling for multiple comparisons, students in the treatment group did not exhibit statistically significant lower levels of high risk or delinquent behavior relative to students in the control group.
- Subgroup Findings:
    - Subgroup analyses were conducted to examine both impacts across groups, and also whether impacts were statistically significant within subgroups.
    - For boys, the impact of the program on future orientation was positive and statistically significant.
    - Impacts on girls were statistically significantly different from boys for the self-reported Pro-social Behaviors and Scholastic Efficacy and School Bonding scales. Additionally,
        - For boys only, the impact of the program on the Pro-social Behaviors scale was negative and statistically significant.
        - For girls only, the impact on the Scholastic Efficacy and School Bonding scale was positive and statistically significant.
    - The impact on truancy did not statistically significantly differ across age groups. However,
        - The impact on truancy was positive and statistically significant for students below age 12, but not for students aged 12 and older.

- There were no statistically significant findings for subgroups defined by family structure, academic risk, or baseline delinquency, either within or across subgroups.

- Correlations Between Site-Level Characteristics and Impacts:

    - A series of correlational analyses were conducted to examine whether there were statistically significant relationships between site-level characteristics and impacts. Given that these analyses were conducted outside of the randomized experimental design, the findings reported here cannot be used to draw causal inferences.

    - At the site level, we found that ongoing mentoring support, as measured by the frequency of mentor/supervisor meetings, was statistically significantly associated with site-level impacts across a range of outcome measures from all three impact domains. However, site-level impacts were negatively associated with mentor support for all of these outcomes. For the other eight program characteristics examined in our correlational site-level analyses, there were several statistically significant findings but they were inconsistent with regards to being positively or negatively associated with site-level impacts on student outcomes.

## 4.2    Summarizing Baseline Student Outcomes and Characteristics

This section summarizes the characteristics of the sample of students who participated in the Impact Study. As shown in Exhibit 4.1, the study sample constituted a fairly diverse group of students with the following key characteristics:

- Gender:   Boys constituted 47 percent of the sample versus 53 percent for girls.

- Age:  The average age of the sample was 11.2 years, with 30 percent of the students aged 12 or older.

- Race/Ethnicity:  The sample was predominantly of minority status; 22 percent of the students were white.  Forty-one percent of the student sample was black or African American, and an additional 31 percent was Hispanic.

- School lunch eligibility status:  The sample was of low socio-economic status; 86 percent of the sample was eligible for either free or reduced price lunch.

- Family structure:  Fifty-six percent of the sample came from two-parent households.

- Prior experience in mentoring:  Twenty-six percent of the sample reported receiving mentoring in the prior school year.

As a preliminary step in our analyses, we compared baseline student characteristics and outcome measures across treatment and control groups to determine whether the treatment and control groups were statistically equivalent prior to treatment.[58] Overall differences in observable characteristics[59]

---

[58]   Random assignment ensures balance across treatment and control groups within programs when using observation weights to adjust for varying probabilities of assignment to treatment, as described earlier in this section. However, it does *not* ensure balance across treatment and control groups as a whole. Furthermore, some differences between treatment and control groups even within programs may occur through chance alone; pooling treatment and control groups across sites may mask any individual within-program treatment-control differences that do occur.

between the two groups of students were assessed using a regression-based model with a pooled treatment effect and individual site-level dummy indicators. Results of this analysis are reported in Exhibit 4.1.

---

[59]  Similarities between these groups along observable characteristics do not imply that they are also identical along unobservable characteristics.

**Exhibit 4.1: Comparison of Baseline Characteristics Between Treatment and Control Groups**

| Descriptive Variables | Entire Sample | | Treatment Group | | Control Group | | Difference | Significance (p-value) |
|---|---|---|---|---|---|---|---|---|
| | Unadjusted Mean | Standard Deviation[a] | Unadjusted Mean | Standard Deviation | Unadjusted Mean | Standard Deviation | | |
| Two parents (Percent) | 56.36 | | 56.09 | | 56.63 | | -0.39 | 0.84 |
| Male (Percent) | 47.37 | | 47.21 | | 47.53 | | -0.21 | 0.91 |
| 12 years old or older (Percent) | 29.82 | | 30.01 | | 29.63 | | 0.34 | 0.82 |
| White (Percent) | 21.76 | | 21.97 | | 21.55 | | 0.47 | 0.71 |
| Free or reduced price lunch (Percent) | 85.60 | | 84.15 | | 87.05 | | -2.92* | 0.02 |
| Involved in a mentoring program in the previous school year (Percent) | 25.97 | | 26.01 | | 25.92 | | 0.18 | 0.92 |
| Number of students | 2573 | | 1272 | | 1301 | | | |
| Percent missing data | ≤2 | | ≤3 | | ≤2 | | | |
| **Outcome Variables** | | | | | | | | |
| *Interpersonal Relationships, Personal Responsibility, and Community Involvement* | | | | | | | | |
| Pro-social Behaviors | 2.87 | 0.75 | 2.86 | 0.76 | 2.87 | 0.73 | -0.01 | 0.57 |
| Number of students | 2573 | | 1272 | | 1301 | | | |
| Percent missing data | ≤4 | | ≤4 | | ≤4 | | | |
| *Academic Outcomes* | | | | | | | | |
| *Self-Reported Outcome* | | | | | | | | |
| Scholastic Efficacy and School Bonding | 3.15 | 0.78 | 3.14 | 0.79 | 3.16 | 0.76 | -0.02 | 0.45 |
| Future Orientation | 3.82 | 0.60 | 3.83 | 0.55 | 3.81 | 0.64 | 0.02 | 0.15 |
| Number of students | 2573 | | 1272 | | 1301 | | | |
| Percent missing data | ≤4 | | ≤5 | | ≤4 | | | |
| *School-Reported Outcome* | | | | | | | | |
| Overall Absenteeism (Percent) | 5.07 | 8.92 | 5.12 | 9.47 | 5.03 | 8.33 | 0.12 | 0.66 |
| Grades (Range 1–5) | | | | | | | | |
| Math | 3.50 | 1.83 | 3.50 | 1.84 | 3.50 | 1.83 | 0.00 | 0.94 |
| English Language Arts | 3.75 | 1.75 | 3.75 | 1.77 | 3.75 | 1.73 | 0.03 | 0.62 |
| Science | 3.89 | 2.14 | 3.87 | 2.22 | 3.90 | 2.07 | 0.03 | 0.58 |
| Social Studies | 3.79 | 1.98 | 3.76 | 2.01 | 3.82 | 1.95 | -0.04 | 0.51 |

**Exhibit 4.1: Comparison of Baseline Characteristics Between Treatment and Control Groups**

| Descriptive Variables | Entire Sample | | Treatment Group | | Control Group | | Difference | Significance (p-value) |
|---|---|---|---|---|---|---|---|---|
| | Unadjusted Mean | Standard Deviation[a] | Unadjusted Mean | Standard Deviation | Unadjusted Mean | Standard Deviation | | |
| State Assessment Tests | | | | | | | | |
|   Math—Percent Proficient | 51.79 | | 52.42 | | 51.16 | | 3.70 | 0.11 |
|   Reading/ELA—Percent Proficient | 50.68 | | 51.02 | | 50.34 | | 1.85 | 0.41 |
|     Number of students | 2573 | | 1272 | | 1301 | | | |
|     Percent missing data | ≤35 | | ≤36 | | ≤34 | | | |
| ***Delinquent Behavior Outcome*** | | | | | | | | |
| *Self-Reported Outcome* | | | | | | | | |
|   Misconduct | 3.30 | 0.80 | 3.29 | 0.80 | 3.30 | 0.80 | -0.01 | 0.73 |
|   Delinquency | 3.89 | 0.33 | 3.90 | 0.33 | 3.89 | 0.32 | 0.00 | 0.66 |
|     Number of students | 2573 | | 1272 | | 1301 | | | |
|     Percent missing data | ≤4 | | ≤4 | | ≤3 | | | |
| *School-Reported Outcome* | | | | | | | | |
|   Truancy—Unexcused Absence Rate (Percent)[b] | 2.43 | 7.52 | 2.55 | 8.45 | 2.30 | 6.52 | 0.29 | 0.26 |
|   Misconduct | | | | | | | | |
|     Percent committing any infraction | 15.46 | | 14.92 | | 16.00 | | -1.07 | 0.49 |
|     Percent committing repeated infractions (2+) | 8.76 | | 8.42 | | 9.11 | | -0.78 | 0.52 |
|   Delinquency | | | | | | | | |
|     Percent committing any infraction | 13.00 | | 12.42 | | 13.59 | | -1.42 | 0.33 |
|     Percent committing repeated infractions (2+) | 5.27 | | 5.72 | | 4.81 | | 0.98 | 0.30 |
|     Number of students | 2573 | | 1272 | | 1301 | | | |
|     Percent missing data | ≤29 | | ≤28 | | ≤29 | | | |

[a]    Standard Deviations are only reported for Means or Mean Percents.

[b]    Based on 27 sites that reported unexcused absences and total days enrolled

Entire sample: Missing data ≤44; Treatment Group: Missing data ≤45; Control Group: Missing data ≤43

\*    p-value (of adjusted difference in means) < 0.05, two-tailed test.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

*If differences between groups on characteristics not already included in our model specification had been detected, we would have added those covariates to our model.* However, across 23 individual comparisons, only one statistically significant difference (with at least 95 percent confidence) across groups was found: a higher proportion of students in the treatment group were eligible to receive free or reduced-price school lunches than in the control group. Note that we would expect to find one statistically significant finding by chance when conducting this many comparisons. Moreover, this covariate was already included in the original specification, so the model was not altered.

## 4.3   Overall Impacts of the Student Mentoring Program

This section reports overall impact findings organized by three research questions. As explained in Chapter 1, these three questions are tied to the legislative intent of the program funding and to the activities supported by the funded programs, as laid out in the logic model (see Exhibit 1.1). For each research question, we report impact estimates for one or more individual outcome measures. The treatment impact on each of these outcome measures is estimated by the difference in regression-adjusted means between the treatment and control groups averaged across sites.

The exhibits in this chapter present a number of statistics related to each reported impact, including the following:

- Unadjusted group means for both treatment and control group students (aggregated across sites, using site-level weights proportional to the inverse of the sample size for each site);

- Standard deviations for outcomes measured on a continuous scale;

- Regression-adjusted treatment-control group differences (not necessarily equal to the difference between the unadjusted group means);

- P-values to test difference between treatment and control groups at 0.05 significance level;

- Benjamini-Hochberg corrected critical values to account for multiple comparisons;

- Estimated effect sizes:

  – For outcomes measured on a continuous scale, effect sizes are expressed in terms of standard deviation units (based on the pooled standard deviation of the two groups).[60] That is, an effect size equal to one would imply that treatment was associated with a change in the outcome of one standard deviation.
  – For binary outcomes, effect sizes are expressed in terms of odds-ratios, which are equal to the odds of "success" in the treatment divided by the odds of "success" in the control group.  For example, for our "Math Proficiency" outcome, the odds ratio is equal to the odds that treatment group students were proficient in math at the end of the sample period, divided by the odds that control group students were proficient in math at the end of the sample period.  An odds ratio equal to one means that the control and treatment groups are equally likely to have success.  On the other hand,

---

[60]   Using effect sizes allows one to more easily compare results from outcomes using different scales of measurement.

an odds ratio of two means that the odds of success in the treatment group were twice as high as the control group, and an odds ratio of one half means that the odds of success in the control group were twice as high as in the treatment group;

- Estimates of the upper bound on the percentage of missing data for each group of outcomes, based on the number of students with valid data from each respective data source.

Estimates of standard errors and 95 percent confidence intervals for all estimates are presented in Appendix F.

**1. What is the impact of ED school-based mentoring programs on students' interpersonal relationships with adults, personal responsibility, and community involvement?**

One goal of ED school-based mentoring programs is to foster improved positive behaviors and interpersonal relationships in students through mentoring. As outlined in Chapter 2 and elaborated upon in Appendix C, we developed a single composite scale, named Pro-social Behaviors, as the single outcome measure for this impact domain. This composite scale ranges from 1 to 4 and has a mean (unadjusted) value in the sample of 2.79. (See Appendix C for a list of the individual items included in this outcome measure.)

The estimated impact on the Pro-social Behaviors scale is reported in Exhibit 4.2. We did not find a statistically significant difference with respect to this outcome in the spring between students in the treatment group and students in the control group.[61]

**Exhibit 4.2**

**Estimated Impact on Interpersonal Relationships, Personal Responsibility, and Community Involvement**

| | Unadjusted Mean Outcome | | | | Estimated Impact | | |
| | Treatment Group | | Control Group | | Regression Adjusted T-C Group Difference[a] | p-value to Test Difference | Estimated Effect Size |
| Self-Reported Outcome (Scale Score: Range 1–4) | Mean | Standard Deviation | Mean | Standard Deviation | | | |
|---|---|---|---|---|---|---|---|
| Pro-social Behaviors | 2.79 | 0.81 | 2.80 | 0.80 | -0.01 | 0.68 | -0.01 |
| Number of Students | 1163 | | 1197 | | | | |
| Percent Missing Data | ≤3% | | ≤3% | | | | |

a   Regression Adjusted T-C Group Difference will not necessarily be equal to the difference between the Unadjusted Mean Outcomes.

*   p-value (of adjusted difference in means) < 0.05, two-tailed test.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007.

---

[61]   Estimated impacts for the four individual outcome measures based on the original Student Survey scales in this impact domain (Peer Relationships, Parental Relationships, Relationships with Other Adults, and Personal Responsibility and Community Involvement) are reported in Appendix D, Exhibit D.2. There were no statistically significant differences in the original form outcome measures between the treatment and control groups after controlling for multiple comparisons. Additionally, we caution the reader to note that, although these measures are more directly comparable to scales used in prior studies, many do not meet standard minimal criteria for internal reliability in this population, as reported in Appendix C.

**2.    What is the impact of ED school-based mentoring programs on students' school engagement (e.g., attendance, positive attitude towards school) and academic achievement?**

We measured school engagement and academic achievement using five types of outcome measures:

- Scholastic Efficacy and School Bonding scale (Student Survey)
- Future Orientation scale (Student Survey)
- Absenteeism rate (student records)
- Math, English language arts, science and social studies grades (student records)
- Math and reading/ELA proficiency (student records)

Exhibit 4.3 shows estimated impacts for the nine school engagement and academic achievement outcome measures used.

Scores on the Scholastic Efficacy and School Bonding scale (range 1–4, unadjusted sample mean 3.04) did not statistically significantly differ between treatment and control groups following the intervention.[62] Though the regression-adjusted mean for the Future Orientation scale (range 1–4, unadjusted sample mean 3.82) was higher among treatment students than among controls, with a p-value of 0.04, the effect was not statistically significant after controlling for multiple comparisons.

Mean grades in math, English language arts, science, and social studies were not statistically significantly different across treatment and control groups. Similarly, we found no statistically significant impact of student mentoring on state performance assessment test scores measured by the percent of students receiving a score of "proficient" or better as defined by each state or school district on math and reading/ELA assessment test scores.

Finally, absenteeism rates were approximately half a percentage point (.46) lower in the treatment group than in the control group at the end of the study year, with a p-value of 0.04. However, this impact was not statistically significant after controlling for multiple comparisons.

---

[62]    Appendix D, Exhibit D.3 shows impact estimates for the two original Student Survey scales which together comprise the composite Scholastic Efficacy and School Bonding outcome. School Bonding did not differ significantly between the treatment and control groups. Although Scholastic Efficacy was higher in the treatment group, with a p-value of 0.02, this difference was not statistically significant after controlling for multiple comparisons. Additionally we caution the reader to note that, although these measures are more directly comparable to scales used in prior studies, they do not meet standard criteria for internal reliability in our sample, as reported in Appendix C.

**Exhibit 4.3: Estimated Impact on Academic Outcomes**

| Self-Reported Outcome (Scale Score: Range 1-4) | Unadjusted Mean Outcome | | | | Estimated Impact | | | |
|---|---|---|---|---|---|---|---|---|
| | Treatment Group | | Control Group | | Regression Adjusted T-C Group Difference[b] | p-value to Test Difference | BH-Corrected Critical Value[c] | Estimated Effect Size |
| | Mean | Standard Deviation[a] | Mean | Standard Deviation | | | | |
| Scholastic Efficacy and School Bonding | 3.06 | 0.80 | 3.03 | 0.85 | 0.04 | 0.08 | 0.02 | 0.07 |
| Future Orientation | 3.85 | 0.54 | 3.80 | 0.63 | 0.03* | 0.04 | 0.01 | 0.08 |
|     Number of Students | 1163 | | 1197 | | | | | |
|     Percent Missing Data | ≤3% | | ≤3% | | | | | |
| **School-Reported Outcome** | | | | | | | | |
| *Overall Absenteeism Rate (all absences as percent of total days enrolled)* [e] | 5.03 | 7.71 | 5.49 | 9.63 | -0.46* | 0.04 | 0.01 | -0.09[d] |
|     Number of Students | 1163 | | 1197 | | | | | |
|     Percent Missing Data | 15% | | 18% | | | | | |
| *Grades (Range 1–5)* [f] | | | | | | | | |
| Math | 3.19 | 1.70 | 3.23 | 1.67 | -0.05 | 0.23 | 0.02 | -0.05 |
| English Language Arts | 3.57 | 1.78 | 3.61 | 1.69 | -0.04 | 0.40 | 0.03 | -0.04 |
| Science | 3.52 | 1.87 | 3.55 | 1.86 | -0.03 | 0.48 | 0.04 | -0.03 |
| Social Studies | 3.53 | 1.92 | 3.56 | 1.83 | -0.01 | 0.78 | 0.05 | -0.01 |
|     Number of Students | 1163 | | 1197 | | | | | |
|     Percent Missing Data | ≤35% | | ≤33% | | | | | |
| *State Assessment Tests* | | | | | | | | |
| Math—Percent Proficient | 45.69 | | 47.10 | | -1.53 | 0.41 | 0.04 | 0.94[f] |
| Reading/ELA—Percent Proficient | 49.40 | | 50.76 | | -1.67 | 0.37 | 0.03 | 0.94[f] |
|     Number of Students | 1163 | | 1197 | | | | | |
|     Percent Missing Data | ≤25% | | ≤20% | | | | | |

[a]   Standard Deviations are only reported for Means or Mean Percents.

[b]   Regression Adjusted T–C Group Difference will not necessarily be equal to the difference between the Unadjusted Mean Outcomes. Regression adjustment provides more statistically precise impact estimates than a simple difference in means.

[c]   Based on Benjamini-Hochberg test; figure shown provides the critical value that the "p-value to Test Difference" in the preceding column must be less than in order for the "Regression Adjusted T-C Group Difference" to be statistically significant after controlling for multiple tests.

[d]   Lower rates for Overall Absenteeism indicate more positive outcomes.

[e]   Odds-ratio

[f]   Higher scores indicate higher grades; See Appendix F for further explanation of how these scores were derived.

\*   p-value (of adjusted difference in means) < 0.05, two-tailed test.

\+   p-value (of adjusted difference in means) < BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**3.    What is the impact of ED school-based mentoring programs on students' high-risk or delinquent behavior?**

Another goal of the Student Mentoring program is to prevent or change negative behaviors, such as gang activity, substance abuse, and general misconduct/delinquency, among participating students. As described in greater detail in Chapter 2, to assess the influence of ED school-based mentoring programs on these negative behaviors, we analyzed measures of both self-reported student misbehavior (from the Student Survey; see Appendix C, Exhibit C.2 for items included in self-reported Misconduct and Delinquency scales) and school-reported misbehavior (from school records; see Appendix C, Exhibit C.5 for a detailed description of the school-reported misconduct and delinquent behavior measures).

Note that school-reported infractions represent a *detection* rate rather than a *commission* rate for misbehavior; that is; the student's misbehavior must both occur and be detected by the school and reported in school records. Consequently, these measures do not reflect undetected types of misbehavior not reported to school administrators. If the intervention differentially affected these unreported behaviors, we would be unable to detect that effect in the school record data. While student self-reports of misbehavior are not subject to this limitation (i.e., they are asked about both serious and less serious risk behaviors), they do rely on the student's memory, self-assessment and honesty. Collecting and analyzing both types of measures rather than just one or the other provides a more comprehensive view of high risk and delinquent behaviors.  Also, it should be noted that several of the outcomes measured in this domain (e.g., gang involvement, alcohol and drug use) represented low incidence behaviors for the students in this study, in particular for the subset of students under age 12.

Exhibit 4.4 shows estimated impacts for the seven outcome measures in this domain. We found no statistically significant impacts on high risk or delinquent behavior after controlling for multiple comparisons. Specifically, analysis of the outcomes for treatment and control groups on the self-reported Misconduct (range 1–4, unadjusted sample mean 3.20) and Delinquency scales (range 1–4, unadjusted sample mean 3.86) from the Student Survey[63] shows no statistically significant difference between groups at the end of the study year on either measure.[64] Similarly, there were no statistically significant differences between treatment and control groups on the five misconduct and delinquency measures based on student records. One of these—Truancy, as measured by the unexcused absence rate—was lower among treatment group students, with a p-value of 0.02, but this difference was not statistically significant after adjusting for multiple comparisons.

---

[63]    Note that these Student Survey measures are coded so that a higher score is associated with *lower* levels of student-reported misbehavior.

[64]    Estimated impacts for the five outcome measures based on the original Student Survey question groupings in this impact domain (delinquent/problem behaviors, gang activity, tobacco use, alcohol use, and drug use) are reported in Appendix D, Exhibit D.4. No statistically significant differences in these outcomes between treatment and control groups were detected.

**Exhibit 4.4**

**Estimated Impact on Delinquent Behaviors and Participation in Harmful Activities**

| Self-Reported Behavioral Outcome[a] (Scale Score: Range 1- 4) | Unadjusted Mean Outcome | | | | Estimated Impact | | | |
|---|---|---|---|---|---|---|---|---|
| | Treatment Group | | Control Group | | Regression Adjusted T-C Group Difference[c] | p-value to Test Difference | BH-Corrected Critical Value[d] | Estimated Effect Size |
| | Mean | Standard Deviation[b] | Mean | Standard Deviation | | | | |
| Misconduct | 3.20 | .086 | 3.20 | .085 | 0.00 | 0.95 | 0.05 | 0.00 |
| Delinquency | 3.87 | .036 | 3.85 | .40 | 0.01 | 0.28 | 0.02 | 0.04 |
| Number of Students | 1163 | | 1197 | | | | | |
| Percent Missing Data | ≤3% | | ≤4% | | | | | |
| **School-Reported Behavioral Outcome**[e] | | | | | | | | |
| ***Truancy (unexcused absences as percent of total days enrolled)*** [f] | 2.04 | 4.8 | 2.47 | 6.91 | -0.45* | 0.02 | 0.01 | -0.14 |
| ***Misconduct*** | | | | | | | | |
| Percent committing any infraction | 25.00 | | 22.91 | | 2.56 | 0.13 | 0.01 | 1.59[g] |
| Percent committing repeated infractions (2+) | 14.21 | | 15.63 | | -0.98 | 0.48 | 0.04 | 0.93[g] |
| ***Delinquency*** | | | | | | | | |
| Percent committing any infraction | 18.13 | | 20.03 | | -1.51 | 0.35 | 0.03 | 0.91[g] |
| Percent committing repeated infractions (2+) | 8.64 | | 9.13 | | -0.56 | 0.65 | 0.04 | 0.93[g] |
| Number of Students | 1163 | | 1197 | | | | | |
| Percent Missing Data | ≤22% | | ≤23% | | | | | |

[a]  Higher scores on the Self-Reported Misconduct and Delinquency scales indicate more positive outcomes.

[b]  Standard Deviations are only reported for Means or Mean Percents.

[c]  Regression Adjusted T-C Group Difference will not necessarily be equal to the difference between the Unadjusted Mean Outcomes.

[d]  Based on Benjamini-Hochberg test.

[e]  Lower percents of the school-reported Truancy, Misconduct, and Delinquency items indicate more positive outcomes.

[f]  Based on 27 sites that reported unexcused absences and total days enrolled.

[g]  Odds-ratio.

Treatment Group: Missing data ≤38%; Control Group:  Missing data ≤36%

*   p-value (of adjusted difference in means) < 0.05, two-tailed test.
+   p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing.

*Source:*  Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

## 4.4 Estimation of Subgroup Effects

We also examined the outcome data for any differences in the impact of ED school-based student mentoring programs between subgroups of students.[65] Given the diversity of our sample, we hypothesized that even if we found no statistically significant measurable impacts for our sample as a whole, there might be differences in the magnitude of impacts for different subgroups of students. To test this hypothesis, we first divided the sample into two subgroups for each of the following five characteristics of interest:

- Gender (boys versus girls),

- Age (students below the age of 12 versus students 12 years and older),

- Family structure (two-parent households versus all others),

- Academic risk (below academic proficiency in either reading/ELA or math at baseline versus proficient in both),[66] and

- Baseline delinquency (self-reported delinquent behaviors at baseline versus no self-reported delinquent behaviors at baseline).[67]

These subgroups were chosen based on association with differences in mentoring impacts found in other studies. For instance, Herrera et al. (2007) found differences in impacts of school-based mentoring by gender and age, and Lee and Cramond (1999) found that students from single-parent families responded more favorably to formal mentoring. Students at higher academic or delinquency risk have also been shown to respond differentially to mentoring. Data from an earlier BBBS study (Grossman, et al., 1999) showed that students with the lowest achievement levels, highest levels of absenteeism and least family support made the most gains in attendance, school performance, and drug avoidance compared to students with moderate or high levels of achievement and family support. To reinforce this point, the meta-analysis of 59 program by DuBois et al. (2002) reported the largest effects obtained from mentoring among youth with both multiple individual and environmental risk factors.[68]

We obtained impact estimates for each of the selected subgroups using the same approach as in the main analysis: within each subgroup, we obtained impact estimates for each individual site using the

---

[65] We chose to examine several factors in subgroup analyses as opposed to including them as interaction terms in the main impact analyses.

[66] Sixty percent of the student sample was defined to be at academic risk under this definition.

[67] One-fourth of the sample (25 percent) was at risk for delinquency, defined by presence of self-reported delinquent behaviors at baseline, including stealing, gang activity, possession of a weapon, and alcohol or drug use.

[68] For example, DuBois et al., (2002) cites larger effect sizes for youth from low socioeconomic backgrounds. In our study, we did not test for low-income status as a moderator of program impacts given the high proportion of low-income students (i.e., 86 percent of students eligible for free/reduced lunch) However, we tested for differences in subgroup impacts for other factors strongly correlated with low-income status such as at-risk for delinquency, at risk for academic failure, and family structure.

same methodology as for the full sample, and then computed an average treatment effect for each subgroup by calculating the average of the site-level impacts, weighted proportionally to the size of the treatment and control groups in each site. We then performed a t-test to identify any statistically significant differences in impacts between each paired set of subgroups – for example, to test whether the estimated impact of school-based student mentoring on boys was different from the impact on girls in our sample. These tests of statistical differences, both within and between subgroups of students, were adjusted for multiple comparisons, using the BH correction, similar to our approach with the overall impacts.

Exhibits 4.5–4.19 display detailed impact estimates by subgroup. In the remainder of this section we discuss statistically significant findings in detail. All impact estimates referred to as "statistically significant" below were statistically significant *at the .05 significance level controlling for multiple comparisons.*[69] It should be noted, as discussed in greater detail in Chapter 2, that the statistical power of our study to detect effects in subgroups is lower than the statistical power to detect effects in the full sample, particularly for the smallest subgroups. Thus, any lack of statistically significant findings may be in part a function of limited power for the subgroup analyses.

**Differences in Impacts by Gender**

*Interpersonal Relationships, Personal Responsibility, and Community Involvement (Pro-social Behavior)*
Estimates of the impact of school-based mentoring programs on Pro-social Behaviors by gender subgroup appear in Exhibit 4.5.[70] Treatment group boys, but not treatment group girls, reported statistically significantly lower scores on the Pro-social Behaviors scale from the Student Survey compared to their control group peers. The difference in impacts between boys and girls was also statistically significant.[71]

*Academic Outcomes*
Impacts on academic outcomes by gender subgroup are reported in Exhibit 4.6. We found that the impact of Student Mentoring Programs on the self-reported Scholastic Efficacy and School Bonding scale was positive and statistically significant for girls, but not for boys. The difference in impacts on

---

[69] It should be noted that for these exhibits, all p-values and BH critical values refer to tests of statistical significance of differences in impact between subgroups. These statistics for within-subgroup impacts are not provided in the exhibits, but are available from the authors upon request.

[70] Note that, because there is only one outcome measure in this impact domain, it was not necessary to perform the Benjamini-Hochberg adjustment for multiple comparisons for this outcome.

[71] Estimated impacts by gender subgroup for the four individual outcome measures based on the original Student Survey scales in this impact domain (Peer Relationships, Parental Relationships, Relationships with Other Adults, and Personal Responsibility and Community Involvement) are reported in Appendix D, Exhibit D.5. There were no statistically significant differences in these outcomes between the treatment and control groups after controlling for multiple comparisons. Additionally, we caution the reader to note that, although these measures are more directly comparable to scales used in prior studies, many do not meet standard minimal criteria for internal reliability, as reported in Appendix C.

Scholastic Efficacy and School Bonding by gender was also statistically significant.[72] There was a statistically significant positive impact of the Student Mentoring Programs on the Future Orientation scale in our study for boys, but not for girls, the opposite of the finding for Scholastic Efficacy and School Bonding. However, the difference in impacts on Future Orientation between boys and girls was not statistically significant after controlling for multiple comparisons. For all other academic outcomes, neither impacts on boys or girls, nor differences in impacts between boys and girls, were statistically significant, mirroring the findings in the full sample.

### Delinquent and Harmful Behaviors

Exhibit 4.7 reports impact estimates for delinquent and harmful behaviors for boys and girls. There were no statistically significant differences for boys or girls on misconduct and the gender difference in impacts was not statistically significant.

---

[72] Appendix D, Exhibit D.6 reports separate estimates for the Scholastic Efficacy and School Bonding scales by gender subgroup. These results are similar in magnitude and statistical significance to the results on the composite Scholastic Efficacy and School Bonding measure.

**Exhibit 4.5: Subgroup Findings by Gender: Interpersonal Relationships, Personal Responsibility, and Community Involvement**

| Self-Reported Outcome (Range 1-4) | Unadjusted Mean Outcome | | | | | | | | Estimated Impacts | | | |
| | Boys | | | | Girls | | | | | | | |
| | Treatment | | Control | | Treatment | | Control | | Estimated Impact on Boys[a] | Estimated Impact on Girls[a] | Difference in Impacts | p-value to test Difference |
| | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | | | | |
| Pro-social Behaviors | 2.71 | 0.83 | 2.78 | 0.80 | 2.86 | 0.78 | 2.81 | 0.80 | -0.06* | 0.04 | -0.10* | 0.01 |
| Number of students | 542 | | 573 | | 621 | | 624 | | | | | |
| Percent missing data | 2% | | 2% | | 2% | | 4% | | | | | |

[a]  Estimated Impacts on Boys and Girls will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

*  p-value (of adjusted difference in means) < 0.05, two-tailed test.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007.

**Exhibit 4.6: Subgroup Findings by Gender: Academic Outcomes**

| | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | | |
| | Boys | | | | Girls | | | | | | | | BH-Corrected Critical Value[c] |
| | Treatment | | Control | | Treatment | | Control | | Estimated Impact on Boys[b] | Estimated Impact on Girls[b] | Difference in Impacts | p-value to test Difference | |
| Self-Reported Outcome (Range 1–4) | Mean | Standard Deviation[a] | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Scholastic Efficacy and School Bonding | 2.96 | | 3.00 | | 3.15 | | 3.05 | | -0.03 | 0.10*+ | -0.12*+ | 0.00 | 0.01 |
| Future Orientation | 3.83 | | 3.76 | | 3.86 | | 3.84 | | 0.07*+ | 0.00 | 0.07* | 0.03 | 0.01 |
| Number of students | 542 | | 573 | | 621 | | 624 | | | | | | |
| Percent missing data | ≤3% | | ≤3% | | ≤3% | | ≤4% | | | | | | |
| **School-Reported Outcome** | | | | | | | | | | | | | |
| *Overall Absenteeism Rate (Percent)* [d] | 5.06 | 7.35 | 5.08 | 8.24 | 5.01 | 8.02 | 5.86 | 10.75 | -0.12 | -0.82* | 0.69 | 0.13 | 0.02 |
| Number of students | 542 | | 573 | | 621 | | 624 | | | | | | |
| Percent missing data | 13% | | 16% | | 16% | | 19% | | | | | | |
| *Grades (Range 1–5)* [e] | | | | | | | | | | | | | |
| Math | 3.16 | 1.79 | 3.19 | 1.71 | 3.29 | 1.74 | 3.34 | 1.78 | -0.07 | -0.04 | -0.03 | 0.74 | 0.04 |
| English Language Arts | 3.54 | 1.82 | 3.56 | 1.78 | 3.83 | 2.17 | 3.92 | 2.00 | -0.08 | 0.05 | -0.13 | 0.16 | 0.03 |
| Science | 3.47 | 2.05 | 3.49 | 2.11 | 3.78 | 2.12 | 3.83 | 2.07 | -0.04 | -0.03 | -0.01 | 0.88 | 0.05 |
| Social Studies | 3.43 | 1.97 | 3.37 | 1.98 | 3.77 | 2.11 | 3.91 | 2.01 | 0.01 | -0.05 | 0.06 | 0.53 | 0.03 |
| Number of students | 542 | | 573 | | 621 | | 624 | | | | | | |
| Percent missing data | ≤39% | | ≤34% | | ≤33% | | ≤35% | | | | | | |
| *State Assessment Tests* | | | | | | | | | | | | | |
| Math—Percent Proficient | 47.44 | | 47.61 | | 44.09 | | 46.60 | | -2.38 | -1.17 | -1.21 | 0.75 | 0.04 |
| Reading/ELA—Percent Proficient | 50.09 | | 48.57 | | 48.78 | | 52.87 | | 1.61 | -4.91 | 6.52 | 0.08 | 0.02 |
| Number of students | 542 | | 573 | | 621 | | 624 | | | | | | |
| Percent missing data | ≤23% | | ≤18% | | ≤25% | | ≤22% | | | | | | |

[a]  Standard Deviations are only reported for Means or Mean Percents.

[b]  Estimated Impacts on Boys and Girls will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

[c]  Based on Benjamini-Hochberg test

[d]  Lower rates for Overall Absenteeism indicate more positive outcomes.

[e]  Higher scores indicate higher grades; See Appendix C for further explanation of how these scores were derived.

\*  p-value (of adjusted difference in means) < 0.05, two-tailed test.

+  p-value (of adjusted difference in means) < BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing..

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit 4.7: Subgroup Findings by Gender: Delinquent Behaviors and Participation in Harmful Activities**

| | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | | |
| | Boys | | | | Girls | | | | | | | | |
| | Treatment | | Control | | Treatment | | Control | | | | | | |
| Self-Reported Outcome (Range 1-4)[a] | Mean | Standard Deviation[b] | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | Estimated Impact on Boys[c] | Estimated Impact on Girls[c] | Difference in Impacts | p-value to test Difference | BH-Corrected Critical Value[d] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Misconduct | 3.17 | 0.86 | 3.16 | 0.91 | 3.23 | 0.85 | 3.24 | 0.79 | 0.02 | -0.01 | 0.04 | 0.38 | 0.03 |
| Delinquency | 3.83 | 0.41 | 3.81 | 0.47 | 3.90 | 0.29 | 3.89 | 0.30 | 0.01 | 0.01 | -0.01 | 0.73 | 0.05 |
| Number of students | 542 | | 573 | | 621 | | 624 | | | | | | |
| Percent missing data | ≤3% | | ≤3% | | ≤2% | | ≤4% | | | | | | |
| **School-Reported Behavioral Outcome** | | | | | | | | | | | | | |
| *Truancy—Unexcused Absence Rate (Percent)* [e,f] | 2.03 | 4.35 | 2.13 | 5.76 | 2.06 | 5.24 | 2.85 | 7.97 | -0.23 | -0.76* | 0.53 | 0.20 | 0.01 |
| Number of students | 542 | | 573 | | 621 | | 624 | | | | | | |
| Percent missing data | 37% | | 36% | | 47% | | 47% | | | | | | |
| *Misconduct* [f] | | | | | | | | | | | | | |
| Percent committing any infraction | 31.59 | | 25.64 | | 18.87 | | 20.36 | | 6.33* | -1.21 | 7.54* | 0.03 | 0.01 |
| Percent committing repeated infractions (2+) | 18.13 | | 17.63 | | 10.56 | | 13.76 | | 0.83 | -2.55 | 3.38 | 0.23 | 0.02 |
| *Delinquency* [f] | | | | | | | | | | | | | |
| Percent committing any infraction | 22.73 | | 24.02 | | 13.85 | | 16.30 | | 0.16 | -2.55 | 2.71 | 0.41 | 0.04 |
| Percent committing repeated infractions (2+) | 10.18 | | 10.58 | | 7.21 | | 7.77 | | 0.61 | -0.97 | 1.59 | 0.52 | 0.04 |
| Number of students | 542 | | 573 | | 621 | | 624 | | | | | | |
| Percent missing data | ≤19% | | ≤21% | | ≤23% | | ≤23% | | | | | | |

[a]   Higher scores on the Misconduct and Delinquency scales indicate more positive outcomes.

[b]   Standard Deviations are only reported for Means or Mean Percents.

[c]   Estimated Impacts on Boys and Girls will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

[d]   Based on Benjamini-Hochberg test.

[e]   Based on 27 sites that reported unexcused absences and total days enrolled.

[f]   Lower percents of the school-reported Truancy, Misconduct, and Delinquency items indicate more positive outcomes.

*   p-value (of adjusted difference in means) < 0.05, two-tailed test.

+   p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing.

*Source:*  Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Differences in Impacts for Other Subgroups**

- *Age*:  Exhibits 4.8-4.10 show estimated impacts for students below age 12 and for students aged 12 and older.  Truancy rates were statistically significantly lower in the treatment group for younger students, but not for older students.  However, no statistically significant *differences* in impacts between age groups were found for any outcome measure (including truancy) across the three impact domains.

- *Family Structure*:  Exhibits 4.11-4.13 report estimated impacts by family structure for each of our three impact domains: academic outcomes, delinquent behaviors and participation in harmful activities, and interpersonal relationships, personal responsibility, and community involvement. As a whole, these findings mirror our main impact findings, showing no statistically significant impacts or differences in impacts for either the students from two-parent or other household subgroups.

- *Academic Risk*:  Results of our impact analyses for subgroups defined by baseline academic proficiency in reading/ELA and math appear in Exhibits 4.14-4.16. Impacts on proficient and non-proficient students did not statistically significantly differ for outcomes in any of the three impact domains, nor were impacts on the individual proficiency-defined subgroups statistically significant. This matches findings for the study sample as a whole.

- *Baseline Delinquency*:  As shown in Exhibits 4.17-4.19, we found no statistically significant differences in impacts between students who reported delinquent behaviors at baseline and students who did not, nor were impacts on the individual delinquency-defined subgroups statistically significant. This matches findings for the entire study sample.

**Exhibit 4.8: Subgroup Findings by Age: Interpersonal Relationships, Personal Responsibility, and Community Involvement**

| Self-Reported Outcome (Range 1-4) | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | |
| | Students Below Age 12 | | | | Students Aged 12 and Older | | | | | | | |
| | Treatment | | Control | | Treatment | | Control | | Estimated Impact on Younger Students[a] | Estimated Impact on Older Students[a] | Difference in Impacts | p-value to test Difference |
| | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | | | | |
| Pro-social Behaviors | 2.84 | 0.80 | 2.86 | 0.79 | 2.68 | 0.81 | 2.66 | 0.80 | -0.02 | 0.03 | -0.05 | 0.25 |
| Number of students | 826 | | 833 | | 337 | | 364 | | | | | |
| Percent missing data | 2% | | 3% | | 2% | | 2% | | | | | |

[a]  Estimated Impacts on Younger and Older Students will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

*  p-value (of adjusted difference in means) < 0.05, two-tailed test.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007.

**Exhibit 4.9: Subgroup Findings by Age: Academic Outcomes**

| Self-Reported Outcome (Range 1–4) | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | | |
| | Students Below Age 12 | | | | Students Aged 12 and Older | | | | | | | | |
| | Treatment | | Control | | Treatment | | Control | | Estimated Impact on Younger Students[b] | Estimated Impact on Older Students[b] | Difference in Impacts | p-value to test Difference | BH-Corrected Critical Value[c] |
| | Mean | Standard Deviation[a] | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | | | | | |
| Scholastic Efficacy and School Bonding | 3.14 | 0.75 | 3.11 | 0.80 | 2.89 | 0.84 | 2.82 | 0.91 | 0.02 | 0.08* | -0.06 | 0.22 | 0.01 |
| Future Orientation | 3.87 | 0.47 | 3.83 | 0.59 | 3.78 | 0.66 | 3.74 | 0.71 | 0.03 | 0.04 | -0.01 | 0.84 | 0.03 |
| Number of students | 826 | | 833 | | 337 | | 364 | | | | | | |
| Percent missing data | ≤3% | | ≤4% | | ≤2% | | ≤2% | | | | | | |
| **School-Reported Outcome** | | | | | | | | | | | | | |
| *Overall Absenteeism Rate (Percent)* [d] | 4.24 | 5.99 | 4.64 | 7.95 | 6.94 | 10.42 | 7.50 | 12.18 | -0.51* | -0.88 | 0.37 | 0.54 | 0.02 |
| Number of students | 826 | | 833 | | 337 | | 364 | | | | | | |
| Percent missing data | 14% | | 18% | | 17% | | 16% | | | | | | |
| *Grades (Range 1–5)* [e] | | | | | | | | | | | | | |
| Math | 3.37 | 1.71 | 3.43 | 1.65 | 2.91 | 1.86 | 2.89 | 1.75 | -0.07 | -0.03 | -0.04 | 0.71 | 0.02 |
| English Language Arts | 3.85 | 1.72 | 3.88 | 1.66 | 3.15 | 1.95 | 3.16 | 1.78 | -0.04 | -0.05 | 0.01 | 0.89 | 0.04 |
| Science | 3.93 | 1.91 | 4.00 | 1.84 | 3.11 | 2.18 | 2.99 | 1.94 | -0.02 | -0.02 | 0.00 | 0.99 | 0.05 |
| Social Studies | 3.89 | 1.99 | 3.90 | 1.91 | 3.11 | 2.03 | 3.08 | 1.72 | 0.02 | -0.01 | 0.03 | 0.78 | 0.03 |
| Number of students | 826 | | 833 | | 337 | | 364 | | | | | | |
| Percent missing data | ≤37% | | ≤35% | | ≤34% | | ≤32% | | | | | | |
| **State Assessment Tests** | | | | | | | | | | | | | |
| Math—Percent Proficient | 50.57 | | 52.84 | | 34.21 | | 33.21 | | -3.24 | 1.13 | -4.37 | 0.28 | 0.01 |
| Reading/ELA—Percent Proficient | 52.96 | | 55.94 | | 40.90 | | 38.39 | | -1.83 | -1.12 | -0.71 | 0.86 | 0.04 |
| Number of students | 826 | | 833 | | 337 | | 364 | | | | | | |
| Percent missing data | ≤24% | | ≤20% | | ≤26% | | ≤21% | | | | | | |

a  Standard Deviations are only reported for Means or Mean Percents.

b  Estimated Impacts on Younger and Older Students will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

c  Based on Benjamini-Hochberg test.

d  Lower rates for Overall Absenteeism indicate more positive outcomes.

e  Higher scores indicate higher grades; See Appendix C for further explanation of how these scores were derived.

*  p-value (of adjusted difference in means) < 0.05, two-tailed test.

+  p-value (of adjusted difference in means) < BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing..

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit 4.10: Subgroup Findings by Age: Delinquent Behaviors and Participation in Harmful Activities**

| | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Students Below Age 12 | | | | Students Aged 12 and Older | | | | | | | | |
| | Treatment | | Control | | Treatment | | Control | | | | | | |
| Self-Reported Outcome (Range 1-4)[a] | Mean | Standard Deviation[b] | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | Estimated Impact on Younger Students[c] | Estimated Impact on Older Students[c] | Difference in Impacts | p-value to test Difference | BH-Corrected Critical Value[d] |
| Misconduct | 3.28 | 0.81 | 3.27 | 0.83 | 3.02 | 0.90 | 3.05 | 0.86 | 0.01 | -0.04 | 0.06 | 0.23 | 0.03 |
| Delinquency | 3.90 | 0.30 | 3.89 | 0.35 | 3.79 | 0.46 | 3.77 | 0.46 | 0.01 | 0.02 | -0.01 | 0.57 | 0.04 |
| Number of students | 826 | | 833 | | 337 | | 364 | | | | | | |
| Percent missing data | ≤3% | | ≤4% | | ≤1% | | ≤2% | | | | | | |
| **School-Reported Behavioral Outcome** | | | | | | | | | | | | | |
| *Truancy—Unexcused Absence Rate (Percent)* [e,f] | 1.60 | 3.67 | 1.90 | 5.44 | 2.98 | 6.50 | 3.78 | 9.06 | -0.55*+ | -0.84 | 0.30 | 0.56 | 0.04 |
| Number of students | 826 | | 833 | | 337 | | 364 | | | | | | |
| Percent missing data | 43% | | 43% | | 39% | | 38% | | | | | | |
| *Misconduct* [f] | | | | | | | | | | | | | |
| Percent committing any infraction | 22.67 | | 19.49 | | 31.00 | | 31.61 | | 4.11* | -3.03 | 7.14 | 0.06 | 0.01 |
| Percent committing repeated infractions (2+) | 12.05 | | 12.63 | | 19.77 | | 23.27 | | 0.46 | -6.32* | 6.78* | 0.03 | 0.01 |
| *Delinquency* [f] | | | | | | | | | | | | | |
| Percent committing any infraction | 14.75 | | 15.25 | | 26.83 | | 32.18 | | -0.22 | -6.40 | 6.17 | 0.12 | 0.02 |
| Percent committing repeated infractions (2+) | 5.01 | | 5.46 | | 17.99 | | 18.47 | | -0.56 | -1.09 | 0.53 | 0.87 | 0.05 |
| Number of students | 826 | | 833 | | 337 | | 364 | | | | | | |
| Percent missing data | ≤20% | | ≤20% | | ≤24% | | ≤28% | | | | | | |

[a]   Higher scores on the Misconduct and Delinquency scales indicate more positive outcomes.

[b]   Standard Deviations are only reported for Means or Mean Percents.

[c]   Estimated Impacts on Younger and Older Students will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

[d]   Based on Benjamini-Hochberg test.

[e]   Based on 27 sites that reported unexcused absences and total days enrolled.

[f]   Lower percents of the school-reported Truancy, Misconduct, and Delinquency items indicate more positive outcomes.

*   p-value (of adjusted difference in means) < 0.05, two-tailed test.
+   p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing.

*Source:*  Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit 4.11: Subgroup Findings by Family Structure: Interpersonal Relationships, Personal Responsibility, and Community Involvement**

| Self-Reported Outcome (Range 1-4) | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | |
| | Two-Parent Households | | | | Other Households | | | | Estimated Impact on Two-Parent HHs[a] | Estimated Impact on Other HHs[a] | Difference in Impacts | p-value to test Difference |
| | Treatment | | Control | | Treatment | | Control | | | | | |
| | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | | | | |
| Pro-social Behaviors | 2.83 | 0.79 | 2.81 | 0.79 | 2.73 | 0.83 | 2.78 | 0.82 | 0.02 | -0.04 | 0.06 | 0.14 |
| Number of students | 705 | | 735 | | 549 | | 556 | | | | | |
| Percent missing data | 9% | | 8% | | 10% | | 12% | | | | | |

[a]   Estimated Impacts on Two-Parent Households and Other Households will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

*   p-value (of adjusted difference in means) < 0.05, two-tailed test.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007.

**Exhibit 4.12: Subgroup Findings by Family Structure: Academic Outcomes**

| | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | | |
| | Two-Parent Households | | | | Other Households | | | | Estimated Impact on Two-Parent HHs[b] | Estimated Impact on Other HHs[b] | Difference in Impacts | p-value to test Difference | BH-Corrected Critical Value[c] |
| | Treatment | | Control | | Treatment | | Control | | | | | | |
| Self-Reported Outcome (Range 1–4) | Mean | Standard Deviation[a] | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Scholastic Efficacy and School Bonding | 3.10 | 0.81 | 3.05 | 0.83 | 3.02 | 0.78 | 2.99 | 0.88 | 0.03 | 0.05 | -0.02 | 0.61 | 0.02 |
| Future Orientation | 3.85 | 0.56 | 3.80 | 0.62 | 3.83 | 0.52 | 3.81 | 0.65 | 0.03 | 0.03 | 0.01 | 0.87 | 0.04 |
| Number of students | 705 | | 735 | | 549 | | 556 | | | | | | |
| Percent missing data | ≤9% | | ≤8% | | ≤11% | | ≤13% | | | | | | |
| **School-Reported Outcome** | | | | | | | | | | | | | |
| *Overall Absenteeism Rate (Percent)* [d] | 4.37 | 6.34 | 4.58 | 8.01 | 5.75 | 8.83 | 6.52 | 10.93 | -0.19 | -0.65 | 0.46 | 0.33 | 0.01 |
| Number of students | 705 | | 735 | | 549 | | 556 | | | | | | |
| Percent missing data | 22% | | 24% | | 21% | | 23% | | | | | | |
| *Grades (Range 1–5)* [e] | | | | | | | | | | | | | |
| Math | 3.32 | 1.86 | 3.33 | 1.82 | 3.30 | 1.76 | 3.25 | 1.76 | -0.09 | -0.01 | -0.09 | 0.34 | 0.01 |
| English Language Arts | 3.71 | 1.91 | 3.70 | 1.85 | 3.61 | 1.92 | 3.60 | 1.67 | -0.03 | -0.02 | -0.01 | 0.88 | 0.04 |
| Science | 3.80 | 2.33 | 3.72 | 2.29 | 3.67 | 2.25 | 3.61 | 1.83 | -0.04 | -0.05 | 0.01 | 0.92 | 0.05 |
| Social Studies | 3.82 | 2.28 | 3.71 | 2.23 | 3.55 | 2.22 | 3.62 | 2.06 | 0.04 | 0.02 | 0.02 | 0.83 | 0.03 |
| Number of students | 705 | | 735 | | 549 | | 556 | | | | | | |
| Percent missing data | ≤42% | | ≤41% | | ≤41% | | ≤37% | | | | | | |
| *State Assessment Tests* | | | | | | | | | | | | | |
| Math—Percent Proficient | 46.70 | | 50.69 | | 44.29 | | 42.22 | | -2.38 | -1.17 | -1.21 | 0.75 | 0.03 |
| Reading/ELA—Percent Proficient | 48.81 | | 50.39 | | 50.31 | | 51.12 | | 1.61 | -4.91 | 6.52 | 0.08 | 0.01 |
| Number of students | 705 | | 735 | | 549 | | 556 | | | | | | |
| Percent missing data | ≤230% | | ≤26% | | ≤30% | | ≤27% | | | | | | |

[a]  Standard Deviations are only reported for Means or Mean Percents.

[b]  Estimated Impacts on Two-Parent Households and Other Households will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

[c]  Based on Benjamini-Hochberg test

[d]  Lower rates for Overall Absenteeism indicate more positive outcomes.

[e]  Higher scores indicate higher grades; See Appendix C for further explanation of how these scores were derived.

\*  p-value (of adjusted difference in means) < 0.05, two-tailed test.

+  p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing..

*Source:*  Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit 4.13: Subgroup Findings by Family Structure: Delinquent Behaviors and Participation in Harmful Activities**

| Self-Reported Outcome (Range 1-4)[a] | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Two-Parent Households | | | | Other Households | | | | Estimated Impact on Two-Parent HHs[c] | Estimated Impact on Other HHs[c] | Difference in Impacts | p-value to test Difference | BH-Corrected Critical Value[d] |
| | Treatment | | Control | | Treatment | | Control | | | | | | |
| | Mean | Standard Deviation[b] | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | | | | | |
| Misconduct | 3.25 | 0.87 | 3.23 | 0.85 | 3.14 | 0.84 | 3.16 | 0.85 | -0.01 | 0.00 | -0.01 | 0.88 | 0.05 |
| Delinquency | 3.88 | 0.36 | 3.85 | 0.41 | 3.86 | 0.36 | 3.85 | 0.38 | 0.01 | 0.00 | 0.01 | 0.67 | 0.04 |
| Number of students | 705 | | 735 | | 549 | | 556 | | | | | | |
| Percent missing data | ≤9% | | ≤9% | | ≤10% | | ≤12% | | | | | | |
| **School-Reported Behavioral Outcome** | | | | | | | | | | | | | |
| *Truancy—Unexcused Absence Rate (Percent )* [e,f] | 1.83 | 4.07 | 1.72 | 4.91 | 2.33 | 5.59 | 3.38 | 8.59 | 0.02 | -0.90* | 0.92* | 0.03 | 0.01 |
| Number of students | 705 | | 735 | | 549 | | 556 | | | | | | |
| Percent missing data | 47% | | 48% | | 46% | | 44% | | | | | | |
| *Misconduct* [f] | | | | | | | | | | | | | |
| Percent committing any infraction | 21.41 | | 19.67 | | 29.42 | | 27.11 | | 1.76 | 2.83 | -1.07 | 0.76 | 0.04 |
| Percent committing repeated infractions (2+) | 10.68 | | 13.78 | | 18.31 | | 17.98 | | -2.16 | 0.30 | -2.46 | 0.39 | 0.02 |
| *Delinquency* [f] | | | | | | | | | | | | | |
| Percent committing any infraction | 13.97 | | 17.60 | | 23.06 | | 23.40 | | -3.56 | -0.06 | -3.50 | 0.30 | 0.01 |
| Percent committing repeated infractions (2+) | 6.17 | | 7.60 | | 11.57 | | 11.18 | | -1.92 | 0.15 | -2.08 | 0.41 | 0.03 |
| Number of students | 705 | | 735 | | 549 | | 556 | | | | | | |
| Percent missing data | ≤30% | | ≤31% | | ≤25% | | ≤25% | | | | | | |

[a]    Higher scores on the Misconduct and Delinquency scales indicate more positive outcomes.

[b]    Standard Deviations are only reported for Means or Mean Percents.

[c]    Estimated Impacts on Two-Parent Households and Other Households will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

[d]    Based on Benjamini-Hochberg test.

[e]    Based on 27 sites that reported unexcused absences and total days enrolled.

[f]    Lower percents of the school-reported Truancy, Misconduct, and Delinquency items indicate more positive outcomes.

\*    p-value (of adjusted difference in means) < 0.05, two-tailed test.

\+    p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit 4.14: Subgroup Findings by Academic Risk: Interpersonal Relationships, Personal Responsibility, and Community Involvement**

| | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | |
| | Proficient Students | | | | Not Proficient Students | | | | | | | |
| | Treatment | | Control | | Treatment | | Control | | Estimated Impact on Proficient Students[a] | Estimated Impact on Not Proficient Students[a] | Difference in Impacts | p-value to test Difference |
| Self-Reported Outcome (Range 1-4) | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Pro-social Behaviors | 2.82 | 0.83 | 2.80 | 0.79 | 2.77 | 0.79 | 2.79 | 0.80 | -0.01 | -0.01 | 0.00 | 0.97 |
| Number of students | 373 | | 398 | | 550 | | 586 | | | | | |
| Percent missing data | 2% | | 3% | | 2% | | 3% | | | | | |

[a] Estimated Impacts on Proficient and Not Proficient students will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

* p-value (of adjusted difference in means) < 0.05, two-tailed test.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007.

**Exhibit 4.15: Subgroup Findings by Academic Risk: Academic Outcomes**

| | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | | |
| | Proficient Students | | | | Not Proficient Students | | | | | | | | |
| | Treatment | | Control | | Treatment | | Control | | | | | | |
| Self-Reported Outcome (Range 1–4) | Mean | Standard Deviation[a] | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | Estimated Impact on Proficient Students[b] | Estimated Impact on Not Proficient Students[b] | Difference in Impacts | p-value to test Difference | BH-Corrected Critical Value[c] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Scholastic Efficacy and School Bonding | 3.11 | 0.79 | 3.06 | 0.84 | 3.03 | 0.79 | 2.99 | 0.86 | 0.01 | 0.05 | -0.04 | 0.44 | 0.02 |
| Future Orientation | 3.88 | 0.44 | 3.83 | 0.60 | 3.83 | 0.58 | 3.79 | 0.67 | 0.03 | 0.04 | -0.01 | 0.73 | 0.03 |
| Number of students | 373 | | 398 | | 550 | | 586 | | | | | | |
| Percent missing data | ≤2% | | ≤3% | | ≤4% | | ≤4% | | | | | | |
| **School-Reported Outcome** | | | | | | | | | | | | | |
| *Overall Absenteeism Rate (Percent)* [d] | 4.16 | 6.20 | 4.91 | 8.50 | 5.19 | 7.55 | 5.81 | 9.81 | -0.60 | -0.72* | 0.13 | 0.79 | 0.04 |
| Number of students | 373 | | 398 | | 550 | | 586 | | | | | | |
| Percent missing data | 6% | | 10% | | 10% | | 14% | | | | | | |
| *Grades (Range 1–5* [e] | | | | | | | | | | | | | |
| Math | 4.04 | 1.93 | 4.12 | 1.99 | 3.18 | 1.84 | 3.12 | 1.72 | -0.03 | -0.02 | 0.00 | 0.99 | 0.05 |
| English Language Arts | 4.53 | 2.60 | 4.57 | 2.52 | 3.61 | 1.82 | 3.63 | 1.72 | 0.01 | -0.07 | 0.09 | 0.39 | 0.01 |
| Science | 4.46 | 2.45 | 4.61 | 2.54 | 3.62 | 2.21 | 3.60 | 2.12 | -0.08 | -0.04 | -0.05 | 0.66 | 0.03 |
| Social Studies | 4.48 | 2.81 | 4.61 | 2.69 | 3.43 | 1.97 | 3.37 | 1.98 | 0.07 | 0.00 | 0.07 | 0.51 | 0.02 |
| Number of students | 373 | | 398 | | 550 | | 586 | | | | | | |
| Percent missing data | ≤34% | | ≤32% | | ≤32% | | ≤31% | | | | | | |
| *State Assessment Tests* | | | | | | | | | | | | | |
| Math—Percent Proficient | 75.04 | | 78.58 | | 26.19 | | 25.36 | | -3.35 | 0.33 | -3.68 | 0.33 | 0.01 |
| Reading/ELA—Percent Proficient | 80.50 | | 83.40 | | 28.83 | | 28.46 | | -1.75 | -0.53 | -1.23 | 0.74 | 0.04 |
| Number of students | 373 | | 398 | | 550 | | 586 | | | | | | |
| Percent missing data | ≤5% | | ≤2% | | ≤4% | | ≤4% | | | | | | |

[a]  Standard Deviations are only reported for Means or Mean Percents.

[b]  Estimated Impacts on Proficient and Not Proficient students will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

[c]  Based on Benjamini-Hochberg test.

[d]  Lower rates for Overall Absenteeism indicate more positive outcomes.

[e]  Higher scores indicate higher grades; See Appendix C for further explanation of how these scores were derived.

\*  p-value (of adjusted difference in means) < 0.05, two-tailed test.

+  p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing..

*Source:*  Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit 4.16: Subgroup Findings by Academic Risk: Delinquent Behaviors and Participation in Harmful Activities**

| Self-Reported Outcome (Range 1-4)[a] | Unadjusted Mean Outcome | | | | | | | | Estimated Impact on Proficient Students[c] | Estimated Impact on Not Proficient Students[c] | Difference in Impacts | p-value to test Difference | BH-Corrected Critical Value[d] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Proficient Students | | | | Not Proficient Students | | | | | | | | |
| | Treatment | | Control | | Treatment | | Control | | | | | | |
| | Mean | Standard Deviation[b] | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | | | | | |
| Misconduct | 3.25 | 0.82 | 3.23 | 0.85 | 3.16 | 0.88 | 3.18 | 0.87 | 0.02 | -0.02 | 0.04 | 0.40 | 0.01 |
| Delinquency | 3.88 | 0.34 | 3.87 | 0.39 | 3.86 | 0.36 | 3.84 | 0.40 | 0.01 | 0.01 | 0.00 | 0.85 | 0.04 |
| Number of students | 373 | | 398 | | 550 | | 586 | | | | | | |
| Percent missing data | ≤2% | | ≤2% | | ≤3% | | ≤4% | | | | | | |
| **School-Reported Behavioral Outcome** | | | | | | | | | | | | | |
| *Truancy—Unexcused Absence Rate (Percent)* [e,f] | 1.24 | 3.53 | 1.72 | 4.88 | 2.42 | 4.92 | 2.98 | 8.13 | -0.38 | -0.65* | 0.26 | 0.51 | 0.02 |
| Number of students | 373 | | 398 | | 550 | | 586 | | | | | | |
| Percent missing data | 39% | | 40% | | 36% | | 38% | | | | | | |
| *Misconduct* [f] | | | | | | | | | | | | | |
| Percent committing any infraction | 21.02 | | 19.70 | | 25.94 | | 25.77 | | 0.15 | 1.76 | -1.61 | 0.67 | 0.03 |
| Percent committing repeated infractions (2+) | 12.35 | | 13.76 | | 15.11 | | 17.93 | | -2.13 | -1.67 | -0.46 | 0.89 | 0.06 |
| *Delinquency* [f] | | | | | | | | | | | | | |
| Percent committing any infraction | 17.15 | | 17.86 | | 18.41 | | 22.48 | | 1.77 | -4.24 | 6.00 | 0.1 | 0.01 |
| Percent committing repeated infractions (2+) | 6.51 | | 8.84 | | 8.85 | | 10.36 | | -2.31 | -1.75 | -0.57 | 0.84 | 0.04 |
| Number of students | 373 | | 398 | | 550 | | 586 | | | | | | |
| Percent missing data | ≤27% | | ≤30% | | ≤13% | | ≤15% | | | | | | |

[a]   Higher scores on the Misconduct and Delinquency scales indicate more positive outcomes.

[b]   Standard Deviations are only reported for Means or Mean Percents.

[c]   Estimated Impacts on Proficient and Not Proficient students will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

[d]   Based on Benjamini-Hochberg test.

[e]   Based on 27 sites that reported unexcused absences and total days enrolled.

[f]   Lower percents of the school-reported Truancy, Misconduct, and Delinquency items indicate more positive outcomes.

\*   p-value (of adjusted difference in means) < 0.05, two-tailed test.

+   p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing.

*Source:*  Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit 4.17: Subgroup Findings by Baseline Delinquency: Interpersonal Relationships, Personal Responsibility, and Community Involvement**

| Self-Reported Outcome (Range 1-4) | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | |
| | Any Delinquency | | | | No Delinquency | | | | Estimated Impact on Students Reporting Any Delinquency[a] | Estimated Impact on Students Reporting No Delinquency[a] | Difference in Impacts | p-value to test Difference |
| | Treatment | | Control | | Treatment | | Control | | | | | |
| | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | | | | |
| Pro-social Behaviors | 2.64 | 0.81 | 2.64 | 0.85 | 2.84 | 0.80 | 2.86 | 0.77 | -0.02 | -0.01 | -0.01 | 0.78 |
| Number of students | 277 | | 310 | | 886 | | 887 | | | | | |
| Percent missing data | 3% | | 2% | | 2% | | 3% | | | | | |

[a] Estimated Impacts on Any Delinquency and No Delinquency students will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

* p-value (of adjusted difference in means) < 0.05, two-tailed test.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007.

**Exhibit 4.18: Subgroup Findings by Baseline Delinquency: Academic Outcomes**

| Self-Reported Outcome (Range 1–4) | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Any Delinquency | | | | No Delinquency | | | | Estimated Impact on Students Reporting Any Delinquency[b] | Estimated Impact on Students Reporting No Delinquency[b] | Difference in Impacts | p-value to test Difference | BH-Corrected Critical Value[c] |
| | Treatment | | Control | | Treatment | | Control | | | | | | |
| | Mean | Standard Deviation[a] | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | | | | | |
| Scholastic Efficacy and School Bonding | 2.92 | 0.79 | 2.85 | 0.90 | 3.11 | 0.79 | 3.09 | 0.82 | 0.07 | 0.03 | 0.04 | 0.41 | 0.03 |
| Future Orientation | 3.76 | 0.69 | 3.67 | 0.81 | 3.87 | 0.48 | 3.84 | 0.55 | 0.05 | 0.02 | 0.03 | 0.46 | 0.04 |
| Number of students | 277 | | 310 | | 886 | | 887 | | | | | | |
| Percent missing data | ≤4% | | ≤5% | | ≤3% | | ≤3% | | | | | | |
| **School-Reported Outcome** | | | | | | | | | | | | | |
| *Overall Absenteeism Rate (Percent)* [d] | 6.44 | 9.47 | 6.00 | 11.37 | 4.56 | 6.92 | 5.30 | 8.87 | -0.13 | -0.61* | 0.49 | 0.43 | 0.03 |
| Number of students | 244 | | 272 | | 747 | | 715 | | | | | | |
| Percent missing data | 12% | | 12% | | 16% | | 19% | | | | | | |
| *Grades (Range 1–5)* [e] | | | | | | | | | | | | | |
| Math | 3.09 | 1.74 | 3.06 | 1.73 | 3.24 | 1.70 | 3.31 | 1.67 | 0.00 | -0.08 | 0.07 | 0.49 | 0.04 |
| English Language Arts | 3.54 | 2.42 | 3.50 | 2.10 | 3.76 | 1.83 | 3.81 | 1.74 | 0.06 | -0.08 | 0.14 | 0.19 | 0.01 |
| Science | 3.32 | 2.11 | 3.29 | 1.93 | 3.73 | 1.96 | 3.74 | 1.97 | 0.03 | -0.06 | 0.09 | 0.41 | 0.03 |
| Social Studies | 3.30 | 1.99 | 3.25 | 1.86 | 3.74 | 2.02 | 3.79 | 1.93 | 0.09 | -0.07 | 0.16 | 0.16 | 0.01 |
| Number of students | 199 | | 241 | | 639 | | 649 | | | | | | |
| Percent missing data | ≤42% | | ≤36% | | ≤34% | | ≤34% | | | | | | |
| *State Assessment Tests* | | | | | | | | | | | | | |
| Math—Percent Proficient | 46.19 | | 43.20 | | 45.55 | | 48.40 | | 2.38 | -3.17 | 5.55 | 0.21 | 0.02 |
| Reading/ELA—Percent Proficient | 51.83 | | 48.46 | | 48.68 | | 51.52 | | -0.03 | -2.19 | 2.17 | 0.63 | 0.05 |
| Number of students | 205 | | 247 | | 696 | | 726 | | | | | | |
| Percent missing data | ≤27% | | ≤22% | | ≤23% | | ≤20% | | | | | | |

[a] Standard Deviations are only reported for Means or Mean Percents.

[b] Estimated Impacts on Any Delinquency and No Delinquency students will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

[c] Based on Benjamini-Hochberg test

[d] Lower rates for Overall Absenteeism indicate more positive outcomes.

[e] Higher scores indicate higher grades; See Appendix C for further explanation of how these scores were derived.

\* p-value (of adjusted difference in means) < 0.05, two-tailed test.

+ p-value (of adjusted difference in means) < BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing..

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit 4.19: Subgroup Findings by Baseline Delinquency: Delinquent Behaviors and Participation in Harmful Activities**

| Self-Reported Outcome (Range 1-4)[a] | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | | |
| | Any Delinquency | | | | No Delinquency | | | | Estimated Impact on Students Reporting Any Delinquency[c] | Estimated Impact on Students Reporting No Delinquency[c] | Difference in Impacts | p-value to test Difference | BH-Corrected Critical Value[d] |
| | Treatment | | Control | | Treatment | | Control | | | | | | |
| | Mean | Standard Deviation[b] | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | | | | | |
| Misconduct | 2.94 | 0.92 | 2.89 | 0.91 | 3.29 | 0.80 | 3.31 | 0.78 | 0.02 | -0.01 | 0.03 | 0.57 | 0.04 |
| Delinquency | 3.67 | 0.50 | 3.64 | 0.52 | 3.93 | 0.25 | 3.92 | 0.28 | 0.03 | 0.00 | 0.02 | 0.41 | 0.02 |
| Number of students | 277 | | 310 | | 886 | | 887 | | | | | | |
| Percent missing data | ≤4% | | ≤5% | | ≤2% | | ≤3% | | | | | | |
| **School-Reported Behavioral Outcome** | 2.69 | 5.36 | 3.14 | 8.37 | 1.81 | 4.56 | 2.20 | 6.18 | | | | | |
| ***Truancy—Unexcused Absence Rate (Percent)*** [e,f] | 2.77 | | 3.18 | | 1.82 | | 2.19 | | -0.53 | -0.41 | -0.12 | 0.81 | 0.04 |
| Number of students | 277 | | 310 | | 886 | | 887 | | | | | | |
| Percent missing data | 38% | | 34% | | 44% | | 44% | | | | | | |
| ***Misconduct*** [f] | | | | | | | | | | | | | |
| Percent committing any infraction | 35.54 | | 28.42 | | 21.58 | | 21.06 | | 4.59 | 1.36 | 3.22 | 0.43 | 0.03 |
| Percent committing repeated infractions (2+) | 22.58 | | 18.48 | | 11.49 | | 14.68 | | 2.72 | -2.32 | 5.04 | 0.16 | 0.01 |
| ***Delinquency*** [f] | | | | | | | | | | | | | |
| Percent committing any infraction | 26.04 | | 32.13 | | 15.56 | | 15.97 | | -4.36 | -0.36 | -3.99 | 0.33 | 0.01 |
| Percent committing repeated infractions (2+) | 14.81 | | 16.12 | | 6.64 | | 6.79 | | -0.53 | -0.15 | -0.38 | 0.91 | 0.05 |
| Number of students | 277 | | 310 | | 886 | | 887 | | | | | | |
| Percent missing data | ≤18% | | ≤25% | | ≤22% | | ≤22% | | | | | | |

[a]   Higher scores on the Misconduct and Delinquency scales indicate more positive outcomes.

[b]   Standard Deviations are only reported for Means or Mean Percents.

[c]   Estimated Impacts on Any Delinquency and No Delinquency students will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

[d]   Based on Benjamini-Hochberg test.

[e]   Based on 27 sites that reported unexcused absences and total days enrolled.

[f]   Lower percents of the school-reported Truancy, Misconduct, and Delinquency items indicate more positive outcomes.

\*   p-value (of adjusted difference in means) < 0.05, two-tailed test.

+   p-value (of adjusted difference in means) < BH-Corrected Critical Value ➜ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing.

*Source:*  Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

## 4.5  Site-Level Characteristics and Impacts

Characteristics of programs and their mentors varied considerably across sites. Although we did not find that the Student Mentoring Programs had statistically significant impacts on student-level outcomes for our sample as a whole, we wished to determine whether some sites or groups of sites could be characterized as more or less successful, and, if so, whether we could identify program characteristics associated with differences in impacts at the site level.[73]

Appendix G, Exhibits G.1-G.17 display site-level impacts and 95 percent confidence intervals for each of seventeen outcome measures. As is evident from these graphs, impact estimates were not generally statistically distinguishable from one another at the site level. A series of F tests for each outcome measure confirmed that impact estimates did not statistically significantly vary across sites for any of the seventeen outcomes. This lack of statistically significant variation in impacts across sites limits the ability to make statistical inferences about possible sources of impact variation based on these data.

Additionally, because sites were not randomly assigned to different levels of implementation, no causal inferences can be made, based on these analyses.  Even in the presence of a properly specified statistical model, correlational relationships cannot be used to infer causality. Also, with regards to the correlational findings from these analyses, program characteristics or aspects of the mentor-student interaction may have evolved the way they did because of mentor or student characteristics, not necessarily due to programmatic requirements of the sites themselves.  In other words, the site-level characteristics used in these analyses are impossible to disentangle from either mentor or student characteristics.  This raises the possibility that the correlations explored here are spurious (in part or as a whole) and hence misleading if interpreted as hypothesis-testing evidence of *why* impacts differ.

In summary, the analyses in the remainder of this chapter examine associations between site-level impacts and a selected set of predictors (which may be correlated with a number of other predictors). Thus, the analyses cannot be used to infer causality and are best characterized as correlational and exploratory in nature.

### Choosing Appropriate Site-Level Covariates

In performing the site-level analysis, there are appreciable sample size limitations. With only 42 site-level observations (or fewer, for those outcome measures with missing data at the site level), it was essential to develop a parsimonious model for testing for any relationship between program characteristics (and contextual factors) and site-level impacts. In choosing the final set of site-level covariates for inclusion in our model, we therefore considered several factors:

---

[73]   Note that the statistical power of these analyses is necessarily constrained by the limited sample size of sites.  Thus, any lack of statistically significant findings may be in part a function of limited power for these analyses.  The reader should also note that the estimates produced from these analyses are based on a relatively small number of observations given the number of parameters in the statistical model. The reader should exercise caution in interpreting the statistical estimates obtained from these models, given that estimates from these types of models are generally more consistent as samples become larger.

1. *Theoretical importance in influencing impacts*: Only variables with a theoretically plausible influence on program impacts were considered for inclusion. We developed a list of covariates that we hypothesized might be associated with differential site-level impacts based on:

    - Our review of the literature on mentoring in general, and school-based mentoring in particular;
    - Our understanding of factors that typically affect the efficacy of youth-service interventions, such as baseline academic and behavioral risk factors of targeted students;
    - Our knowledge of the way in which mentoring programs are administered; and
    - The Student Mentoring Program logic model and authorizing legislation.

    In general, we focused on proximal factors thought to have a direct influence on impacts rather than distal factors, which may be correlated with impacts, but whose influence may be indirect and/or mediated through more proximal causes. For example, one could argue that program experience should be correlated with program effectiveness, because more experienced programs are more likely to have developed or implemented methods found to be more successful, such as ongoing mentor support or pre-match mentor training. In this example, program experience is a distal factor associated with two more proximal factors, ongoing support and pre-match training; only the latter two variables were therefore considered for inclusion in our model.

2. *Site-level variation*: Some variables thought to be important determinants of program impacts did not statistically significantly vary across sites. For example, in the Grantee Survey, 91 percent of grantees reported that academic achievement was a major focus of their programs. (Not surprising, considering that this focus was mandated in the original program legislation.) Consequently, the lack of site-level variation in reported program focus on academic achievement prevented us from analyzing the associations between this factor and site-level impacts. Only variables which were shown via F-tests to statistically significantly vary across sites were therefore considered for inclusion in our model.

3. *Site-level correlations across variables:* We examined correlations among all variables under consideration for inclusion in our analysis. Including explanatory variables that are too highly correlated in a multivariate analysis will lead to a multicollinearity problem, resulting in inflated standard errors and a reduced ability to demonstrate that any of the included variables are statistically significantly related to impacts. We therefore selected only variables that were not highly correlated with other included covariates.

Note that while some variables included in these analyses are based on data on program characteristics provided by grantees (e.g., average hours of pre-match training provided to mentors), others based on aggregate mentor reports may be reactive to the individual needs of students, and therefore more likely to be confounded with student baseline characteristics than are grantee data. Second, two included variables are measures of student rather than program characteristics and are thus more contextual in nature.

The final set of site-level covariates in our analysis grouped by type of measure included:[74]

***Program Delivery (based on pre-intervention activities or characteristics of mentors)***
- Average hours of pre-match training provided to mentors;
- Percent of mentors aged 22 or below;
- Percent of mentor/student matches of the same race/ethnicity;[75]

***Program Delivery (based on aggregated mentor reports post-intervention)***
- Amount of ongoing mentor support (average frequency of mentor-supervisor meetings);[76]
- Frequency of use of activities in mentor/student meetings (e.g., percent of mentors reporting almost always/most of the time either working on homework and/or academic skills with students);
- Percent of mentor/student matches lasting 6 months or longer;
- Average total hours of mentor/student meetings per month.

***Student Characteristics***
- Percent of students with self-reported delinquent behaviors at baseline;
- Percent of students scoring "not proficient" in either math or reading/ELA at baseline;

**Analytic Approach**

We performed a series of ordinary least squares regression analyses. The dependent variable in each specification was the site-level impact estimate.[77] For each of our seventeen outcome variables, the site-level impact estimates were regressed on the nine covariates listed above plus an indicator

---

[74] For those outcomes based on data from the Mentor Survey, we averaged across mentors to produce site-level estimates. An alternative approach would have been to use treatment group students as the unit of analysis, and average the characteristics of mentors assigned to each of these students. In cases whether a single mentor was assigned to mentor more than one student, the former approach would give that mentor equal weight with all other mentors in the analysis, while the latter would count that mentor's characteristics once for each assigned student. Since the outcomes of interest were measured at the student level, the latter approach seems most appropriate; however, since, as described in Chapter 3, a very small proportion of mentors in our sample were assigned to more than one student, we assumed that averaging across mentors instead of across students would have no appreciable influence on our results.

[75] We also measured the percent of same-gender matches. The literature is mixed about the relative efficacy of cross versus same gender matches. Given that many programs decided to match on the basis of both gender and race/ethnicity and the presence of less site-level variation on same gender matching, it was decided to solely focus on the percent of same race/ethnicity matching to make the model more parsimonious.

[76] Frequency of mentor/student meetings was measured on a six point scale, ranging from 1 = "never met" to 6 = "met weekly."

[77] Note that this analytic approach did not take into account dependencies between impact measures in those 10 sites that provided data for two separate study years, as a hierarchical linear modeling framework would have done. Because this was intended as an exploratory analysis, we determined that the small amount of resulting false precision in the site-level estimates was unlikely to appreciably influence results or interpretation.

variable for the share of the control group that received mentoring (from any source) during the outcome period to adjust for potential differential attenuation of impact estimates from site to site.

**Results**

Results of these site-level analyses organized by impact domain are reported in Exhibits 4.20–4.22. In presenting these findings, we again caution the reader not to draw causal inferences given that they are correlational analyses[78].

Statistically significant findings at the 95 percent confidence level are:

- The proportion of mentors aged 22 or younger was negatively associated with impacts on math grades, all other factors held constant (Exhibit 4.21; see also Appendix G, Exhibit G-5).

- The proportion of mentor/student matches of the same race/ethnicity was positively associated with impacts on English language arts grades, all other factors held constant. (Exhibit 4.21; see also Appendix G, Exhibit G-6).

- The frequency of mentor/supervisor meetings was negatively associated with site-level impacts on the Pro-social Behaviors measure from the Student Survey (Exhibit 4.20; see also Appendix G, Exhibit G-1) and on grades in math and social studies, adjusting for the influence of other included covariates (Exhibit 4.21; see also Appendix G, Exhibits G-5 and G-8).

- Mentor/supervisor meeting frequency was also positively associated with site-level impacts on school-reported delinquency from student records (Exhibit 4.22; see also Appendix G, Exhibit G-16).

- The average monthly hours of mentor/student meetings were positively associated with site-level impacts on the Future Orientation measure from the Student Survey (Exhibit 4.21; see also Appendix G, Exhibit G-3), all other included factors held constant, but negatively associated with site-level impacts on grades in math and English language arts (Exhibit 4.21; see also Appendix G, Exhibits G-5 and G-6).

- All other factors held constant, the proportion of students with self-reported delinquent behaviors at baseline was positively associated with site-level impacts on social studies grades, (Exhibit 4.21; see also Appendix G, Exhibit G-8). Similarly, the proportion of students with self-reported delinquent behaviors at baseline was negatively associated with site-level impacts on absenteeism and truancy (Exhibits 4.21 and 4.22; see also Appendix G, Exhibits G-4 and G-13). However, it was also positively associated with

---

[78]  For the purposes of reporting associations between site-level characteristics and impacts, we refer to relationships as "positive" or "negative" in the statistical sense, reflecting the direction of the coefficient. However, in some cases a positive statistical relationship denotes a negative substantive relationship or a negative statistical relationship denotes a positive substantive relationship.

site-level impacts on repeated misconduct from student records, all other factors held constant (Exhibit 4.22; see also Appendix G, Exhibit G-15).

**Exhibit 4.20**

**Site-Level Associations: Relationship between Individual Site Characteristics (Other Measured Characteristics Held Constant) and Student Mentoring Program Impacts on Pro-social Behaviors**

| Program Characteristic | Measure of Association between Characteristic and Program Impact[a] (p-value) |
|---|---|
| Hours of mentor pre-match training/orientation | 0.00 (0.86) |
| Frequency mentor talked w/ supervisor (range 1= never to 6= weekly) | -0.08* (0.02) |
| Almost Always/Mostly worked on relationship building | 0.16 (0.33) |
| Percent of mentors 22 years or younger | 0.06 (0.48) |
| Mentor and student share race/ethnicity | 0.07 (0.54) |
| Any self-reported student delinquency at baseline | -0.12 (0.61) |
| Student not proficient in either math or reading/ELA at baseline | -0.06 (0.62) |
| Percent matches lasting 6+ months | -0.03 (0.76) |
| Average monthly hours mentors met with student | 0.01 (0.28) |
| Percent control group students receiving mentoring | 0.14 (0.51) |

[a]  Coefficients from multivariate regression.

\*  p-value<.05, two-tailed test

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit 4.21**

**Site-Level Associations: Relationship between Individual Site Characteristics (Other Measured Characteristics Held Constant) and Student Mentoring Program Impacts on Academic Outcomes**

| Program Characteristic | Measure of Association between Characteristic and Program Impact[a] (p-value) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Scholastic Efficacy & School Bonding | Future Orientation | Absenteeism Rate | Math Grades | English Language Arts Grades | Science Grades | Social Studies Grades | Math Proficiency | Reading/ELA Proficiency |
| Hours of mentor pre-match training/orientation | 0.00 (0.74) | -0.01 (0.30) | 0.00 (0.15) | -0.03 (0.06) | 0.01 (0.75) | -0.01 (0.79) | -0.02 (0.40) | 0.00 (0.71) | -0.01 (0.29) |
| Frequency mentor talked w/ supervisor (range 1= never to 6= weekly) | -0.01 (0.71) | -0.03 (0.17) | 0.01 (0.10) | -0.11* (0.05) | 0.03 (0.67) | -0.06 (0.41) | -0.22* (0.01) | -0.05 (0.13) | -0.02 (0.55) |
| Almost Always/Mostly worked on academic skills or homework | 0.00 (0.99) | -0.07 (0.53) | -0.01 (0.51) | -0.03 (0.90) | -0.27 (0.47) | -0.45 (0.22) | 0.00 (0.99) | 0.07 (0.69) | 0.01 (0.94) |
| Percent of mentors 22 years or younger | 0.05 (0.57) | -0.03 (0.62) | 0.01 (0.50) | -0.29* (0.04) | 0.23 (0.25) | -0.13 (0.49) | 0.08 (0.71) | 0.06 (0.46) | -0.05 (0.61) |
| Mentor and student share race/ethnicity | -0.01 (0.96) | 0.02 (0.80) | -0.01 (0.35) | 0.29 (0.12) | 0.56* (0.04) | 0.18 (0.48) | 0.06 (0.83) | -0.03 (0.83) | 0.12 (0.39) |
| Any self-reported student delinquent behaviors at baseline | -0.15 (0.57) | 0.04 (0.82) | -0.04* (0.05) | 0.22 (0.54) | 0.09 (0.86) | -0.07 (0.89) | 1.39* (0.02) | 0.37 (0.12) | 0.32 (0.23) |
| Student not proficient in either math or reading/ELA at baseline | 0.11 (0.44) | 0.01 (0.88) | -0.01 (0.48) | 0.08 (0.66) | 0.27 (0.32) | 0.28 (0.29) | 0.02 (0.96) | 0.17 (0.19) | -0.10 (0.49) |
| Percent matches lasting 6+ months | 0.11 (0.38) | 0.13 (0.09) | 0.00 (0.77) | 0.09 (0.63) | -0.26 (0.32) | 0.12 (0.64) | 0.30 (0.31) | 0.02 (0.86) | -0.08 (0.53) |
| Average monthly hours mentors met with student | 0.02 (0.08) | 0.01* (0.04) | 0.00 (0.15) | -0.03* (0.02) | -0.06* (0.00) | -0.01 (0.59) | -0.01 (0.55) | -0.01 (0.40) | -0.01 (0.45) |
| Percent control group students receiving mentoring | 0.12 (0.62) | 0.03 (0.86) | 0.01 (0.75) | 0.01 (0.97) | -0.11 (0.82) | 0.14 (0.77) | 0.13 (0.83) | 0.09 (0.69) | 0.04 (0.87) |

[a]   Coefficients from multivariate regression.

*   p-value<.05, two-tailed test

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit 4.22**

**Site-Level Associations: Relationship between Individual Site Characteristics (Other Measured Characteristics Held Constant) and Student Mentoring Program Impacts on Delinquent Behaviors/Participation in Harmful Activities**

| Program Characteristic | Measure of Association between Characteristic and Program Impact[a] (p-value) | | | | | | |
|---|---|---|---|---|---|---|---|
| | Misconduct (Student Survey) | Delinquency (Student Survey) | Truancy Rate | Any Misconduct (School Records) | Repeated Misconduct (School Records) | Any Delinquency (School Records) | Repeated Delinquency (School Records) |
| Hours of mentor pre-match training/orientation | 0.01 (0.16) | 0.00 (0.82) | 0.07 (0.35) | 0.00 (0.76) | -0.01 (0.31) | 0.01 (0.25) | 0.01 (0.30) |
| Frequency mentor talked w/ supervisor (range 1= never to 6= weekly) | -0.06 (0.06) | -0.02 (0.21) | 0.07 (0.78) | 0.04 (0.37) | 0.03 (0.17) | 0.07* (0.03) | 0.01 (0.69) |
| Almost Always/Mostly worked on risk avoidance | 0.12 (0.42) | 0.12 (0.13) | 0.30 (0.79) | 0.01 (0.94) | -0.03 (0.76) | -0.15 (0.31) | -0.10 (0.34) |
| Percent of mentors 22 years or younger | 0.02 (0.75) | -0.01 (0.81) | 0.27 (0.66) | -0.01 (0.88) | -0.03 (0.60) | 0.09 (0.19) | 0.03 (0.51) |
| Mentor and student share race/ethnicity | 0.04 (0.69) | -0.04 (0.44) | -0.87 (0.28) | 0.00 (0.99) | -0.03 (0.72) | -0.14 (0.20) | -0.03 (0.65) |
| Any self-reported student delinquent behaviors at baseline | 0.07 (0.73) | 0.00 (0.96) | -3.37* (0.02) | 0.15 (0.58) | 0.35* (0.02) | -0.07 (0.70) | -0.05 (0.73) |
| Student not proficient in either math or reading/ELA at baseline | 0.06 (0.61) | -0.08 (0.20) | -0.56 (0.47) | -0.06 (0.76) | -0.11 (0.27) | -0.17 (0.23) | 0.02 (0.81) |
| Percent matches lasting 6+ months | 0.04 (0.68) | -0.01 (0.83) | -0.59 (0.43) | -0.12 (0.44) | -0.12 (0.12) | 0.02 (0.86) | -0.04 (0.63) |
| Average monthly hours mentors met with student | 0.01 (0.14) | 0.01 (0.05) | -0.01 (0.83) | 0.00 (0.84) | 0.00 (0.41) | 0.00 (0.68) | 0.00 (0.80) |
| Percent control group students receiving mentoring | 0.02 (0.93) | -0.05 (0.66) | 2.72 (0.10) | 0.07 (0.78) | 0.15 (0.25) | 0.11 (0.53) | -0.02 (0.90) |

[a]  Coefficients from multivariate regression.

*  p-value<.05, two-tailed test.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

## 4.6   Summary of Impact Analysis

This section summarizes the experimental findings of the Impact Study, as well as the results of our exploratory site-level analyses.  Note that because our sample of grantees was purposively selected, and disproportionately comprised of larger, more experienced, school-based mentoring programs, one must be cautious in extrapolating our findings to the larger population of Student Mentoring Program grantees.

**Overall and Subgroup Impacts**

The findings reported previously in this chapter indicate that, after adjusting for multiple comparisons, ED's Student Mentoring Program did not lead to statistically significant impacts on any of the 17 outcomes in the three outcome domains investigated.

Subgroup analyses did reveal that impacts were somewhat heterogeneous by gender. Impacts on girls were statistically significantly different from impacts on boys for two self-reported scales:  Scholastic Efficacy and School Bonding, and Pro-social Behaviors.  Additionally, for boys only, the impact on self-reported Pro-social Behaviors was negative and statistically significant.[79] In contrast, for girls only, the impacts on Scholastic Efficacy and School Bonding and on the Overall Absenteeism Rate were positive and statistically significant. There was also a statistically significant negative impact on truancy for younger students (below of 12), but not for older students (i.e., indicating that the rate of truancy was lower in the treatment group compared to the control group); however, the differences in impacts between older and younger students were not statistically significant on any of the outcome measures. There were no statistically significant findings for other subgroups defined by family composition, baseline academic non-proficiency, or baseline delinquency, either within or between subgroups.

**Exploratory Analysis of Site Characteristics and Site-Level Impacts**

*Because the lack of statistically significant variation in impacts across sites limits the ability to make statistical inferences about possible sources of impact variation based on these data, and because we do not explicitly control for multiple comparisons in our site-level analyses, these findings must be considered exploratory in nature.  In addition, because these are correlational findings no causal inferences should be drawn from exploratory analyses.*

There were negative associations between program supervision of mentors and site-level impacts on three of the seventeen individual outcome measures: Pro-social Behaviors, math grades, and social studies grades.  There was also a positive association between program supervision of mentors and site-level impacts on delinquency infractions from school records
Relationships of program characteristics to impacts for the other eight characteristics examined in the exploratory site-level analyses were generally inconsistent in direction and statistical significance. For example, the proportion of mentors aged 22 or younger was negatively associated with site-level

---

[79]   The possibility of negative effects of mentoring has some precedent in the mentoring literature. Rhodes (2002) speculates about the possibility of unintentional negative effects of mentoring, particularly in cases where mentoring relationships are disrupted or terminated. In his rigorous experimental evaluation, Karcher (2008) found negative effects of school-based mentoring on cooperation of high-school-aged boys that is consistent with this possibility

impacts on math grades, while the proportion of mentor/student matches of the same race/ethnicity was positively associated with site-level impacts on English language arts grades. The proportion of students with self-reported delinquent behaviors at baseline in each site was positively associated with site-level impacts on social studies grades and repeated misconduct as reported by school records, but negatively associated with site-level impacts on absenteeism and truancy rates.

It is impossible with these data to establish any causality with regards to the findings regarding the largely negative association between student outcomes and both mentor supervision and the number of hours that mentors and student meet. In other words, these findings could indicate that sites with these delivery features are leading to poorer outcome for their students, or students with more problems lead to mentors asking for more help and spending more time with their students. The relationship could also be spurious.

## 4.7    Overall Summary

The Student Mentoring Program is designed to fund grantees to enable them to provide mentoring to at-risk students in grades 4-8. The ultimate goal of the program is to improve student academic and behavioral outcomes through the guidance and encouragement of a volunteer mentor. We measured a total of 17 impacts in the domains of academic achievement and engagement, interpersonal relationships and personal responsibility, and high-risk or delinquent behavior. The main finding of the Impact Study was that there were no statistically significant impacts of the Student Mentoring Program for the sample as a whole on this array of student outcomes. However, there was some scattered evidence that impacts were heterogeneous across types of students. In particular, impacts on girls were statistically significantly different from impacts on boys for two self-reported scales: Scholastic Efficacy and School Bonding, and Pro-social Behaviors. For boys, the impact on Pro-social Behaviors was negative and statistically significant (i.e., treatment group boys had lower Pro-social Behaviors scores). For girls, the impact on Scholastic Efficacy and School Bonding was positive and statistically significant. The impact on truancy was negative and statistically significant for students below age 12 (i.e., younger treatment groups students were less truant than younger control groups students). There were negative associations between program supervision of mentors and site-level impacts on three of the seventeen individual outcome measures—Pro-social Behaviors, grades in math and social studies, and a positive relationship with  the outcome of school-reported delinquency.

Chapter 3 of this report also presented results demonstrating that the Student Mentoring Program represented a fairly low level of intensity in terms of service delivery. Specifically, students received approximately an hour a week of mentoring over a period of six months. Compared to community-based mentoring programs, for example, where students meet with their mentors for 2 or 3 hours per week for 12 months or more (cf. Tierney and Grossman, 2000; Herrera et al., 2007), school-based mentoring represents a more limited opportunity for students to develop enduring, trusting relationships with adult role models. It should be noted, however, that grantees, on average, adhered to the general intents of the legislation and program guidance, while, at the same time being constrained by the limits of the school calendar and the population from which to draw mentors.

Finally, it should be noted that 35 percent of the control group students reported receiving mentoring either from the program or elsewhere in the community. This finding, coupled with the fact that not

all treatment group students met with a mentor, reduced the treatment contrast and may have led to some dilution of the impacts on students compared to expectations.[80]

---

[80] We assume that students seeking other mentoring services in the community typically participated in community-based mentoring programs, which, as previously reported, may represent a different level of intervention in contrast to school-based mentoring.

# References

Aseltine, R., Dupre, M., and Lamlein, P. (2000). Mentoring as a drug prevention strategy: An evaluation of Across Ages. *Adolescent and Family Health 1*(1), 11–20.

Benjamini, Y., and Hochberg, Y. (1995). Controlling for the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological) 57*(1), 289-300.

Bogan, G.A., and Liang, B. (2005). Gender in mentoring relationships. In D.L. DuBois and M.J. Karcher (Eds.), *Handbook of youth mentoring* (pp. 205–217). Thousand Oaks, CA: Sage Publication.

DeVellis, R.F. (2003). *Scale development: Theory and applications* (*2nd edition*). Thousand Oaks, CA: Sage Publications.

DuBois, D.L, Holloway, B.E., Valentine, J.C., and Harris, C. (2002). Effectiveness of mentoring programs for youth: A meta-analytic review. *American Journal of Community Psychology*, *30*(2), 157–197.

DuBois, D.L., and Neville, H. (1997). Youth mentoring: Investigation of relationship characteristics and perceived benefits. *Journal of Community Psychology, 25*(3), 227–234

DuBois D., and Silverthorn, N. (2005). Characteristics of natural mentoring relationships and adolescent adjustment: Evidence from a national study. *The Journal of Primary Prevention*, *26*(2), 69–92.

Eccles, J., and Gootman, J. (2002). *Community programs to promote youth development.* Washington, DC: National Academy Press.

Grossman, J.B., Baldwin, J. and Johnson, A. (1999). Assessing the effectiveness of mentoring programs. In *Contemporary Issues in Mentoring*. Philadelphia: Public/Private Ventures.

Hansen, K. (2005). School-based mentoring and activities during the 2003/2004 school year. Philadelphia, PA: Big Brother/Big Sisters of America.

Harter, S. (1988). *Manual for Self Perception Profile for Adolescents*. University of Denver Press.

Hawkins, J., Guo, J., Hill, K., Battin-Pearson, S., and Abbott, R. (2001). Long term effects of the Seattle social development intervention on school bonding trajectories. *Applied Developmental Science, 5*(4), 225-235.

Herrera, C. (2004). *School-based mentoring: A closer look*. Philadelphia: Public/Private Ventures.

Herrera, C., Grossman, J.B., Kauh, T.J., Feldman, A.F., and McMaken, J. (2007). *Making a difference in schools: The Big Brothers Big Sisters school-based mentoring impact study.* Philadelphia: Public/Private Ventures.

Herrera, C., Sipe, C.L. and McClanahan, W.S. (2000). *Mentoring school-age children: Relationship development in community-based and school-based programs*. Philadelphia: Public/Private Ventures.

Karcher, M.J. (2005). The effects of developmental mentoring and high school mentors' attendance on their younger mentees' self-esteem, social skills, and connectedness. *Psychology in the Schools, 42*(1), 65–77.

Karcher, M.J. (2008). The Study of Mentoring in the Learning Environment (SMILE): A randomized study of the effectiveness of school-based mentoring. *Prevention Science* (Online First [7]). Retrieved February 28, 2008 from http://www.springerlink.com/ content/104965/?Content+Status=Accepted

Lee, J.J., and Cramond, B. (1999). The positive effects of mentoring economically disadvantaged students. *Professional School Counseling, 2*, 172–178.

Nakkula, M.J., and Harris, J.T. (2005). Assessment of mentoring relationships. In D.L. DuBois and M.J. Karcher (Eds.), *Handbook of youth mentoring* (pp. 100–117). Thousand Oaks, CA: Sage Publications.

Portwood, S.G., and Ayers, P.M. (2005). Schools. In D.L. DuBois and M.J. Karcher (Eds.), *Handbook of youth mentoring* (pp. 336–347). Thousand Oaks, CA: Sage Publications.

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. *Journal of Primary Prevention, 26*(2), 129–141.

Rhodes, J.E. (2002). *Stand by me: The risks and rewards of mentoring today's youth.* Cambridge, MA: Harvard University Press.

Rhodes, J., Reddy, R., Roffman, J., and Grossman, J. (2005). Promoting successful youth mentoring relationships: A preliminary screening questionnaire. *Journal of Primary Prevention* 26: 147-168.

Rhodes, J., Grossman, J.B., and Resch, A.L. (2000). Agents of change: Pathways through which mentoring relationships influence adolescents' academic adjustment. *Child Development, 71*(6), 1662–1671.

Schochet, P. (2005). *Statistical power for random assignment evaluations of education programs.* Princeton, NJ: Mathematica Policy Research, Inc.

Sanchez, B., and Colon, Y. (2005). Race, ethnicity and culture in mentoring. In D.L. DuBois and M.J. Karcher (Eds.), *Handbook of youth mentoring* (pp. 191–204). Thousand Oaks, CA: Sage Publication.

Sipe, C.L. (1996). *Mentoring: A synthesis of P/PV's research: 1988–1995.* Philadelphia: Public/Private Ventures.

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

Taylor, A.S., LoSciuto, L., Foz, M., and Sonkowsko, M. (1999). The mentoring factor: Evaluation of the Across Ages intergenerational approach to drug abuse prevention*, Child and Youth Services, 20*(1–2), 77–99.

Thompson, L.A., and Kelly-Vance, L. (2001). The impact of mentoring on academic achievement of at risk youth. *Children and Youth Services Review, 23*(3), 227–242.

Tierney, J.P., and Grossman, J.B. (2000). *Making a difference: An impact study of Big Brothers Big Sisters.* Philadelphia: Public/Private Ventures.

Williams, V.S.L., Jones, L.V. and Tukey, J.W. (1999).  Controlling error in multiple comparisons, with examples from state-to-state differences in educational achievement. *Journal of Educational and Behavioral Statistics, 24*(1), 42-69.

# Appendix A: Sampling Design and Methodology

## Statistical Power

### Background

This section builds upon the brief discussion of statistical power presented in Section 2.4. For our Impact Study, we wished to determine how many grantees and students we needed to recruit in order to detect specified minimum detectable effect sizes for our overall impact analyses. We therefore performed statistical power analysis in order to estimate the required sample size.

A typical standard for statistical power is 0.80; that is, studies typically require sufficient power to detect some minimum specified effect size 80 percent of the time. For the Impact Study, we wished to recruit a sufficient number of grantees and students to detect an effect size of approximately 0.10 with power of 0.80; as discussed in greater detail below, an effect size of this magnitude represents the lower bound on the anticipated effect size for our study, as identified based on the DuBois, Holloway, Valentine, and Harris (2002) meta-analysis of previous mentoring studies. In the remainder of this section, we describe the statistical process used to derive our estimates of anticipated statistical power. Our achieved level of statistical power was very close to our initial goal (see Section 2.4), and was motivated by the statistical derivations provided here.

### Anticipated Model Specification

To perform a power calculation, one must make detailed assumptions about the methodological approach to be used in estimating impacts. When performing our initial power calculations for the Impact Study, we anticipated using ordinary least squares regression to estimate the treatment effect $\beta_{1j}$ for each site. Our power calculations were based on a multi-site randomized trial, where randomization occurred within sites. Under this fixed-effects design, student treatment and control group characteristics are expected to be spread equally across all sites in the study sample. In other words, because we bypassed selection of sites, the connection between students and sites is removed and inferences pertain only to study students in those sites, and any clustering effects between sites are not accounted for (Schochet, 2005).

The anticipated regression was:

$$[A.1] \quad Y_{ij} = \sum_{j=1}^{J} \beta_{1j} T_{ij}\, S_j + \sum_{j=1}^{J} \beta_{2j} S_j + \beta_3\, X_{ij} + \varepsilon_{ij},$$

where,

$Y_{ij}$ is the outcome of interest $Y$ for student $i$ in site $j$,

$T_{ij}$ is the treatment indicator for student $i$ in site $j$ ($T_{ij} = 1$ if student $i$ is assigned to the treatment group; $T_{ij} = 0$ otherwise),

$S_j$ is a program indicator equal to 1 for students randomized at site $j$ and to 0 otherwise ($j = 1...J$),

$\beta_{1j}$ is the estimated average ITT treatment effect for site $j$,

$\beta_{2j}$ is the site-level fixed effect at site $j$ (i.e., the average untreated outcome level of a student at site $j$),

$X_{ij}$ is a vector of student characteristics measured for each student $i$ in site $j$,

$\beta_3$ represents the vector of coefficients indicating how student characteristics affect student outcomes, and

$\varepsilon_{ij}$ represents a random error term for student $i$ in site $j$, independent and identically distributed across students.

Assuming that the error terms are identically and independently distributed, for $K$ total parameters estimated in equation [A.1] and $N$ total students, the estimated variance of the least-squares impact estimate $\hat{\beta}_{1j}$ is

$$[A.2] \quad Var(\hat{\beta}_{1j}) \approx \frac{\hat{\sigma}_j^2}{\sum_{i=1}^{N}(T_{ij} - \overline{T}_j)^2} \quad \text{for each given site } j = [1...J],$$

where

$$[A.3] \quad \hat{\sigma}_j^2 = \sum_{i=1}^{N} \frac{(Y_{ij} - \hat{\beta}_{1j}T_{ij} - \hat{\beta}_{2j} - \hat{\beta}_3 X_{ij})^2}{N - K + 1}.$$

The anticipated ordinary least squares specification would result in $J$ estimates of the treatment effect, one for each of the $J$ sites. To obtain an estimate of the overall treatment effect, we anticipated calculating the weighted average of each of these site-specific estimates, with weights $\omega_j$ set inversely proportional to the sampling variances, $\hat{\sigma}_j^2$.[1] The resulting overall treatment effect estimate would then be given by:

$$[A.4] \quad \hat{\overline{\beta}}_1 = \sum \omega_j \hat{\beta}_{1j},$$

with estimated variance equal to

---

[1]    Application of these weights cancels out the sampling variance term ($\hat{\sigma}_j^2$) in equation [A.2], leaving only the factor of proportionality, which is equal to the numerator term in equation [A.3] when summed across all $J$ sites and all $N$ students.

$$[\text{A.5}] \quad \hat{\sigma}^2 = \sum_{j=1}^{J} \omega_j^2 Var\left(\hat{\beta}_{1j}\right).$$

## Anticipated Effect Size for Overall Impacts

In order to complete our power calculations, in addition to specifying our intended methodological approach, it was also necessary to identify the expected effect size $E$ we anticipated finding in the Impact Study. We grounded the power analysis on the best estimates of student mentoring treatment effect sizes available in the existing literature at the time the analysis was performed. Effect sizes potentially vary by outcome and program characteristics. Because the NCLB legislation authorizing the Student Mentoring Program focused on interpersonal relationships, school performance, and delinquency, we identified likely effect sizes for these types of outcomes.

The most comprehensive source of data on anticipated effect sizes in the context of student mentoring at the time we performed the power analysis was a meta-analysis of 59 mentoring studies conducted by DuBois, et al. (2002). This analysis found an average effect size (under a fixed effects assumption) of 0.14 with a confidence interval of 0.10 - 0.18. We therefore initially assumed an expected effect size of 0.14 in our power analysis. However, we later found that, due to difficulties Student Mentoring Program grantees typically encountered in finding appropriate matches for each student in a timely manner, that the expected amount or "dose" of mentoring, that students actually received from the programs in the current study was less than initially anticipated. We therefore revised our expected effect size assumption to $E = 0.10$, the lower bound of the confidence interval in the DuBois, et al. (2002) meta-analysis.

## Power Calculations for Overall Impacts

The power calculations presented here were performed with respect to the overall impact analyses of the study. Recall that the power of a test statistic is the probability that the null hypothesis $\overline{\beta}_1 = 0$ will be rejected when in reality $\overline{\beta}_1 = \Delta \neq 0$. For the above specification, assuming a two-tailed test at a significance level of $\alpha = .05$, power can be expressed as:

$$[\text{A.6}] \quad POWER = -\int_{-1.96 - \frac{\Delta}{\sigma}}^{1.96 - \frac{\Delta}{\sigma}} \phi(b)db, \text{ where } \phi \text{ denotes the standard normal pdf.}[2]$$

---

[2]   This formula applies exactly when $Y_{ij}$ is a continuous measure with an error term ($\varepsilon_{ij}$ in equation [A.1]) that has a normal distribution. It applies as a close approximation when $Y_{ij}$ is binary, since: (i) estimates of coefficients on binary explanatory variables in linear regression equations, such as the $\beta_{1j}$ coefficients on the $T_{ij}S_j$ explanatory variables in equation [A.1], are essentially sample means on $Y_{ij}$ and differences in sample means on $Y_{ij}$ when derived using least-squares methods; (ii) by the law of large numbers, the asymptotic distribution of any sample mean or difference in sample means is normal, regardless of the distribution of the initial $Y_{ij}$ values for individual students (a Bernoulli distribution in this case, since $Y_{ij}$ is binary); and (iii) our sample sizes of students are large.

We used this formula to perform power calculations by solving for the value of $\Delta$ that yields POWER = 0.80 under a variety of simulated conditions. The parameter $\sigma$ is derived from equations [A.2], [A.3], and [A.5] above.

It is possible to manipulate equation [A.6] to see that the lower the expected variance of the estimated treatment effect, $\hat{\sigma}^2$, the higher the resulting estimate of statistical power. The assumed magnitude of $\hat{\sigma}^2$ in our simulations depended on a number of factors:

- *Sample size*: From equation [A.3], statistical power estimates depended on the assumed number of sites and the number of students within each site from which we would be able to collect data. In general, the larger the number of students and the higher the assumed response rates, the smaller the sampling variance, and the greater the resulting statistical power. However, as seen below, this was also dependent both on the distribution of students across sites and the proportion of students assigned to the treatment group within each site.

- *Proportion assigned to treatment*: From equation [A.2], estimated sampling variance (and thus statistical power) depended on assumptions about the distribution of students between treatment and control groups within each site (i.e., $\overline{T}_j$). A fifty-fifty assignment to treatment (i.e., $\overline{T}_j = 0.5$) minimized sampling variance and thus maximized statistical power in a program; a disproportionate assignment to treatment (in either direction) increased the sampling variance, reducing estimated statistical power. Our power calculations made various assumptions about the number of students per site and how they were distributed between treatment and control groups.

- *Statistical precision*: From equation [A.3], estimated sampling variance (and thus statistical power) varied depending on assumptions about how the introduction of covariates would reduce unexplained variance in outcomes. The lower the degree of unexplained variance, the higher the resulting estimated statistical power. Our power calculations assumed alternatively that covariates explained 25 percent or 50 percent of residual variance (after accounting for treatment) within each program.

- *Distribution of students across sites*: From equation [A.5], estimated sampling variance (and thus statistical power) depended partly on our assumptions about how students were distributed across sites. Sites with more students generally had lower sampling variances for the estimated treatment effect (although the variances also depended on the proportion of students randomly assigned to treatment within each site, as explained above; see equation [A.2]).

Note that the anticipated power calculations we performed were necessarily approximate, based on our prior assumptions about the likely sample size, follow-up data collection response rates, distribution of students across sites and between treatment and control groups, residual variation in outcomes controlling for the covariates and program fixed effects (i.e., the numerator term in equation [A.3] when summed across all $J$ sites as well as all $N$ students), and the exact estimation methodology to be employed. In fact, we could not know the sample size, or how it would be distributed across sites and between treatment and control groups, prior to conducting the experiment.

---

**Target Sample Size for Overall Impacts**

As a result of the power calculations described above, we concluded that a sample size of 2,658 students would be adequate to detect the desired minimum detectable effect size (MDE) of 0.10 with power of 0.80. More specifically, the minimum effect size the study would have an 80 percent chance of detecting as statistically significant ranged from 0.086 to 0.110 for our anticipated sample of 2,658, depending on assumptions about response rates (either 85 or 95 percent) and the explanatory power of the model (covariates explaining either 25 percent or 50 percent of residual variance within each program after accounting for treatment).[3] Comparing this minimum detectable effect to actual effect size findings from previous studies, including the DuBois, et al. (2002) meta-analysis and the BBBSA evaluation (Herrera, 2004), we were confident that, in recruiting this sample size, we would have adequate power to detect reasonable effects for the Impact Study (see discussion in Chapter 2 of the main text).

**Achieved Statistical Power for Overall Impacts**

In fact, as described in Section 2.9, the analytic approach and methodology ultimately employed in our impact analysis differed slightly from the specification we had assumed in our power calculations. First, because sampling rates varied over time or across schools within some sites, our impact analysis introduced observation weights into our regressions to balance the treated and untreated samples equally across time periods and/or schools within each sites. These observation weights added some additional uncertainty to the estimation. However, it is unlikely that weighting greatly changed statistical power relative to our advance calculations.

Second, we eventually chose to assume that treatment effects were heterogeneous across sites, so to estimate the average treatment effect, we weighted site-specific treatment effects by the proportion of students within each site rather than by the inverse of the sampling variance as our power calculations had assumed. Again, this alternative weighting scheme is unlikely to have greatly influenced statistical power, because sampling variances are roughly proportional to the inverse of sample size as discussed in Appendix E.

It was necessary to recruit students in two waves to meet our recruitment targets. The final group of subjects recruited for our study included 2,573 students.[4] As previously displayed in Exhibit 2.1, based on actual sample size, response rates, distribution of students across sites and between treatment and control groups, model specification, and explanatory power of included covariates, our study realized minimum detectable effect sizes associated with overall impacts, ranging from 0.101 to 0.176 across our set of outcome variables.

---

[3] Minimum detectable effect sizes are larger for subgroup analyses due to segmentation of the full sample. (See Section 2.9 for a more detailed discussion of our analytic approach for subgroup analyses.) Under the same assumptions as in the above analysis for the full sample, the minimum detectable effect size rises by 41 percent when estimating impacts on any subset of students that comprises half of the total sample (e.g., boys), and doubles for subsets of students comprising a quarter of the sample.

[4] In total, as previously described in Chapter 2 and in greater detail in the following section, students were recruited from 32 individual sites. Ten programs provided students in both rounds of recruitment and random assignment, for a total of 42 individual groups of students used in our analysis.

## Site Selection and Recruitment Procedures

This section elaborates on the process of how program grantees were selected and recruited into the Impact Study, previously outlined in Chapter 2. Drawing from the pool of 245 grantees representing the total population for the evaluation, members of the study team identified eligible grantees through a review of grant applications and follow-up telephone calls. As explained in Chapter 2, to be selected for the Impact Study, grantees were required to be operational by the fall of the study year. We narrowed this pool further by requiring that programs had to have a sufficient over-subscription of students to support the needs of random assignment. Finally, programs had to indicate a willingness to comply with the needs of the Impact Study with respect to random assignment and data collection. The final pool of grantees meeting these criteria comprised 117 programs. We then rank-ordered sites based on (1) the grantee's estimate of potential demand for services, and (2) indications of the grantee's cooperation with study goals. Top-ranking sites were then targeted for recruitment into our study. The first recruitment phase occurred in Summer-Fall 2005, when 21 grantees, and subsequently 1,329 students, were recruited into the study. The second phase took place in Spring-Fall 2006, when 1,244 additional students were recruited from 21 sites.[5]

Potential grantees were assigned to senior staff recruiters. Recruitment activities included a preliminary letter from ED introducing the study, and a telephone follow-up describing the study and assessing the grantee's interest and capability to participate. Abt staff conducted site visits with all eligible and willing grantees to further explain the study design, develop a Memorandum of Understanding (MOU), discuss any necessary local IRB or school board approval processes, and obtain the program's formal agreement to participate in the study.

Abt staff members were assigned to act as Site Coordinators for each grantee included in the Impact Study to facilitate all communication. In addition, each grantee designated a Site Liaison to work with the Site Coordinator to facilitate communication between the grantee and the study team.

## Identification and Random Assignment of Student Participants for the Impact Study

This section discusses how students were identified and randomly assigned for the Impact Study.

### Identification of Eligible Students

When identifying students for the study, grantees had categorical criteria to determine eligibility, such as grade level or school. Sites also identified appropriate students in a variety of ways, most often asking school staff (such as teachers or counselors from the participating schools) to identify and refer students in need of mentoring to the program.

The Site Liaisons at each program worked with grantee staff to contact the parents of eligible students and obtain signed consent for their children to participate in the study. Common recruitment strategies included sending study consent forms home with students to give to parents to sign and return,

---

[5]    Of the 32 grantees recruited, 10 provided students in both rounds of recruitment and random assignment. In other words, 10 of the 21 grantees recruited in 2006 had previously been recruited (and were part of our sample) in 2005.

advertisements in the local media, and group meetings for parents of eligible students. The study sample comprised all eligible students, though not all students and parents consented to be included in the study.

All grantees also identified "wildcards," or students who were considered in extreme need of mentoring services and were thus allowed to receive mentoring outside of the study. In addition, some programs had state or local requirements to serve specific students, such as children of prisoners. Both wildcard students and students to whom grantees were legally bound to provide services were given mentoring, but were excluded from the study to preserve random assignment.

## Conducting Random Assignment

Specific procedures for random assignment depended upon the timing of individual grantees' recruitment activities and readiness to match students with mentors after random assignment. Sites could either submit multiple student lists during ongoing recruitment of mentees and mentors (i.e., rolling recruitment), or recruit the entire sample and send a final student list to Abt for use in random assignment.

Abt staff conducted the random assignment, matching student IDs with a computerized random sequence of numbers. From these randomly ordered lists, students were sorted into the treatment group, beginning at the top of the list and moving down, selecting as many students as available mentors reported by the program. The lists of students selected for treatment were sent back to programs, so that grantees could begin offering services to these students.[6] This flexible approach to random assignment allowed us to create a randomly ordered list of all eligible applicants, filling program slots with those at the top of the list and assigning the remainder to the control group.

Those students at the bottom portion of the list whose names were not sent to the program were placed into the control group. These students comprised an ordered wait list for slots opening up if students dropped out or new mentors became available within one month of the date that the program began matching students with mentors. In total, approximately two percent of students in the control group moved into the treatment group from the wait list.

Those students who never received a program slot comprised the final control group for the study. Because any segment of a randomly ordered list constitutes a random sample of the overall list, the resulting control group was still well matched to the program group; the only requirement was that slots were offered to applicants in the order their names appeared on the list and that an adequate number of controls remained for the analysis (i.e., at least half of the treatment group sample).

## Maintaining the Integrity of Random Assignment

Any study involving random assignment must monitor program compliance with assignment of participants to avoid contamination of the experimental design. A particularly harmful form of contamination happens when control group individuals "cross over" and receive treatment group services, out of compliance with the randomly-ordered list of students.

---

[6]    For sites with specific requirements for matching mentors with students, these lists were also stratified by school grade level, ethnicity, gender, etc. as necessary.

Although this contamination could not be completely prevented, Site Coordinators minimized crossover by working closely with Site Liaisons to monitor student matching and regularly reminding Site Liaisons that control group students must remain unmatched. Site Coordinators worked with Site Liaisons to check group assignments regularly, and to hold debriefing telephone calls with program directors to review the recruitment and random assignment process, reinforce the importance of maintaining the integrity of random assignment, and discuss appropriate strategies for filling new open program slots from the wait list.[7] Site Liaisons also notified Abt staff if any control group students were matched with a mentor or if there were any other changes in the student sample that affected data collection or analysis, e.g., if a student transferred to a different school.

To preserve random assignment, the study retained drop-outs as members of the treated group. Similarly, wait-list (i.e. control group) students who were not moved to treatment status in order of random assignment were considered control group students, regardless of whether they actually received mentoring from the program as a crossover.

Exhibit A.1 shows the flow of students in the sample, from their original treatment or control group status at the time of initial random assignment to their final status in the study. As described in detail in Section 2.9 of the main text, the model underlying the random assignment and analysis of students was based on an **Intent-To-Treat** framework. In other words, a student's experimental *status* as a treatment or control student, rather than the actual *receipt* of mentoring, served as the measure of treatment. As a result, when a student was informed by a grantee that he or she was allowed to receive services from the program (i.e. that he or she was initially assigned to be a treatment student), then he or she was considered a *treated* student regardless of whether he or she ultimately received mentoring, or the nature of the services actually received. Thus, with the exception of control group students chosen from the wait list in compliance with the randomly-ordered list, who were considered to be valid replacements and therefore re-assigned to treatment group status, all other students in the sample retained their original treatment or control group status in the analysis.

---

[7]  As described above, in the event that programs wished to fill additional slots with students from the control group, they were provided with the first names in sequence from the randomized control group list to fill those slots. In this case, the student(s) from the vacated slot(s) remained in the treatment group. Programs were restricted from this practice if doing so would result in an imbalance of treatment to control group students of more than a 2:1 ratio.

**Exhibit A.1**

**Disposition of Students in the Study Sample**

### ORIGINAL STUDENT STATUS

| TREATMENT GROUP (1292) | | | | CONTROL GROUP (1328) | | |

### FINAL STUDENT STATUS

| **T** Original treatment student (1233) | **R** **Replacement:** student selected from the control group in random order (8) | **D** **Drop:** student refused a mentor or dropped out of program (31) | **Z** **Ineligible** Excluded from random assignment (20) | **C** Original control student (1257) | **Y** **Ineligible** Excluded from random assignment (27) | **X** **Crossover:** student crossed over into the T group OUT OF RA ORDER (44) |

# Selection of the Random Sample of Grantees

Data capturing program characteristics and implementation were collected for each of the 32 purposively selected Impact Study programs, as well as for a stratified random sample of 100 grantees from the pool of 245 eligible programs. The purpose of the random sample was to provide additional descriptive information to ED for program improvement purposes, as well as to determine whether observable characteristics for the Impact Study purposive sample were comparable to those for the universe of ED program grantees.

The rationale for selecting a sample size of 100 grantees was based on the desire to maintain cost-efficiency through dedicating adequate resources to ensure a very high response rate, essential to yielding unbiased sample estimates of population characteristics.[8] Second, we wanted to draw a sample that would provide fairly precise estimates of population grantee characteristics for descriptive purposes as well as for conducting comparisons with the 32 purposively selected programs from the Impact Study. As shown in Exhibit A.4, our sample of 100 programs enjoyed excellent levels of precision.

For the comparison sample of grantees, our goal was to select a random sample of 100 grantees that would be representative of all ED-funded Student Mentoring Program grantees. We therefore stratified programs by auspice (community-based organization, faith-based organization and school district) and by year of funding (2004 or 2005), and then selected a random sample of programs within each stratum. This stratification approach was rooted in the assumptions that 1) programs

---

[8] In fact, our response rate was 100 percent for the administration of the Grantee Survey.

operating under different auspices may have recruited different types of mentors, and 2) programs funded in different years may have represented different levels of experience or stability at the time of the Grantee Survey. Each of these conditions would be expected to result in more homogeneous characteristics of interest in each grouping.

Using stratification in the selection of grantees was beneficial in two ways: 1) it guarded against extreme cases where non-representative draws may have occurred by chance,[9] and 2) it could potentially reduce the variance of overall estimates of program attributes by eliminating variation across strata. To minimize the variance, we stratified the sample in proportion to the number of grantees in each stratum among the 245 funded eligible programs.

**Sample Allocation**

A stratified random sample of 100 grantees was selected. Since we were interested in overall estimates, the best allocation in terms of minimizing the standard errors of the estimates was to allocate the sample in proportion to the number of grantees in each stratum, based on the population of 245 funded programs for which data were available.

Exhibit A.2 shows the distribution of the population of grantees by year and type (school-, community-, or faith-based), and Exhibit A.3 shows the sample allocation.

**Exhibit A.2**

**Distribution of the Population of Grantees**

| Year of Funding | School-Based | Community-Based | Faith-Based | Total |
|---|---|---|---|---|
| Year 1 | 89 | 55 | 10 | 154 63% |
| Year 2 | 50 | 34 | 7 | 91 37% |
| Total | 139 (57%) | 89 (36%) | 17 (7%) | 245 (100%) |

---

[9] This was especially important because some groups of grantees (e.g. faith-based organizations) made up a relatively small proportion of the relevant universe, and could have potentially been left out of a simple random sample drawn without stratification. However, during data analysis and reporting, the community-based and faith-based categories were combined.

**Exhibit A.3**

**Distribution of the Sample by Stratum**

| Year of Funding | School-Based | Community-Based | Faith-Based | Total |
|---|---|---|---|---|
| Year 1 | 36 | 23 | 4 | 63 |
| Year 2 | 20 | 14 | 3 | 37 |
| Total | 56 | 37 | 7 | 100 |

**Selection of the Sample in Each Stratum**

A systematic sample of grantees was drawn in each stratum. We first sorted each stratum's population of programs by region to create an ordered list. We then identified the appropriate sampling interval by dividing the total population of each stratum by the desired number of grantees in that stratum. Finally, we selected grantees from the list, starting at a random point in the first sampling interval, and continuing to select grantees with list placements corresponding to each successive sampling interval end-point.

For example, to select a sample of 20 grantees from the population of 50 in the stratum representing Year 2 school-based programs, the sampling interval was 50/20=2.5. Starting with a randomly-generated number between 1 and 2.5, we then generated 19 more numbers by adding 2.5 successively, rounding the numbers to the closest integer. Grantees whose list placement corresponded to the 20 generated numbers were then selected for the random sample.

Applying this process within each stratum ultimately resulted in a sample representative of the universe of programs by year, auspice, and geographic region. Note that all 32 sites selected for the Impact Study were also contained in the sampling frame used to draw the random representative sample of 100 sites. As seen in the next section of the appendix, this resulted in an overlap between the representative sample and the purposive Impact Study sample, with 12 sites included in both groups.

Exhibit A.4 shows the 95 percent confidence intervals associated with varying population proportions based on drawing a random sample of 100 programs. Note that these confidence intervals have been narrowed by our stratified random sampling process. Stratification increases the precision of the sample estimates if the stratum means are very different.[10] As the exhibit shows, for an observed sample proportion of .30 based on a randomly drawn sample of 100 programs, we can be 95 percent confident that the actual population proportion falls between .23 and .37.[11]

---

[10]   The standard errors underlying these confidence intervals were further adjusted by a small sample finite population correction.

[11]   In other words, if we conducted the random selection of programs 100 times, we would expect the actual population proportion to fall within the estimated confidence interval in 95 instances.

**Exhibit A.4**

**Half-width 95 Percent Confidence Intervals by Population Proportion Based on Population Size, N = 245**

| Population Proportion | Half-Width 95 Percent Confidence Level |
|---|:---:|
| 0.1/.09 | 0.045 |
| 0.2/.08 | 0.060 |
| 0.3/.07 | 0.069 |
| 0.4/.06 | 0.074 |
| 0.5 | 0.075 |

## Testing Differences between Impact Study and Representative Sample Grantees

We wished to determine whether the 32 purposively-selected Impact Study sites differed significantly in terms of observable program characteristics from the full population of 245 Student Mentoring Program grantees funded by ED in 2004 and 2005. Differences in program characteristics of Impact Study sites, if observed, would point to limited generalizability of our Impact Study results for the full universe of Student Mentoring Programs.

Using a difference of means test, we therefore tested the null hypothesis that observable characteristics of the 32 purposively selected Impact Study sites matched those of the full population of 245 Student Mentoring Program grantees funded in 2004 and 2005, as represented by our random sample of 100 sites.[12] For the purposes of this test, we assumed that the observed mean for the 32 Impact Study programs had no sampling variance, since this population was fully observed. In contrast, the mean for the universe of 245 programs was treated as estimated; that is, it had a sampling variance because the estimated mean was based on a random sample of 100 of these grantees. We thus wished to test whether the estimated mean for the 245 programs differed significantly from the known mean for the 32 programs, using a two-tailed test at a significance level of $\alpha=0.05$.

Because the two groups of grantees were not independent (i.e., 12 of the purposively selected Impact Study grantees were also part of the random sample of 100 programs), an independent t test of the difference between means could not be conducted. Nor could a one-sample t test be conducted given that the "population" value (mean of the 100 programs) was only an estimate, and any subsequent test would therefore be statistically inefficient.

Note that the test described above is equivalent to comparing the observed mean for the 32 programs to the estimated mean for the other 213 programs (i.e., the universe of 245 programs minus the 32 purposively selected programs). However, only 88 programs from the random sample of 100 programs were informative for this purpose. The other 12 programs from the sample of 100 programs were included in the purposive Impact Study sample of 32, and so provided no additional

---

[12]   As previously mentioned, although 255 grantees were initially funded, only 245 grantees were eligible for the evaluation.

information.[13] We thus eliminated these 12 programs from our random sample of 100 for the purposes of this comparison.

A sample of 88 represents an appreciable proportion of the 213 programs from the full population not selected for inclusion in the Impact Study. Given this large proportion, and because we wish to make inference about only 213 programs and not some larger population, we applied finite sample adjustments to reduce the variance of the estimated mean for the 213 programs.

Let:

$X_{32}$ = the mean for the 32 Impact Study programs.

$X_{213}$ = the mean for the remaining 213 programs.

$\hat{X}_{213}$ = the estimated mean for the 213 programs based on the sample of 88.

$X_{245}$ = the mean for all 245 programs.

$\hat{X}_{245}$ = the estimated mean for all 245 programs.

Our problem was to compare the mean for the 32 (that is, $X_{32}$, known with certainty) with the mean for the entire population of 245 ($X_{245}$, which we had to estimate). An estimate for the mean of the 245 programs was derived by taking a weighted average of the mean for the 32 purposively selected Impact Study programs and the estimated mean for the other 213 programs.

$$[A.7] \quad \hat{X}_{245} = (32/245) * X_{32} + (213/245) * \hat{X}_{213}$$

The null hypothesis, as stated earlier is:

$$[A.8] \quad X_{32} - X_{245} = 0$$

Substituting the estimated mean $\hat{X}_{245}$ from equation [A.7] for the true population mean $X_{245}$ in equation [A.8], we obtain the following test statistic for the null hypothesis:

$$[A.9] \quad X_{32} - \hat{X}_{245} = X_{32} - (32/245) * X_{32} - (213/245) * \hat{X}_{213}$$
$$= (213/245) * X_{32} - (213/245) * \hat{X}_{213}$$
$$= (213/245) * (X_{32} - \hat{X}_{213})$$

This statistic would be significantly different from zero if the observed mean for the 32 programs differed significantly (based on a two-tailed test at significance level $\alpha=0.05$) from the estimated mean for the 213 programs. All p-values reported for comparisons between the Impact Study sample and the representative random sample of Grantees in Chapter 3 are based on this test statistic.

---

[13]    These 12 programs could have provided an estimate of the mean for the 32 purposively selected Impact Study sites. However, the mean for these 32 sites was already known, so the subsample of 12 provided no additional useful data.

# Appendix B: Survey Instruments

# Grantee Survey

**Please return by**

## Instructions

Please complete all questions, unless you are directed otherwise. When this happens, you will see an arrow with a note that tells you what question to answer like this:

❑₁ Yes  →  *Go to question 2.*
❑₂ No  →  *Go to question 3.*

When answering questions, unless you are directed otherwise, select the *one* answer that best describes your program.

If you have any questions about how to complete the survey, please call the following toll-free number: 1-866-534-9161. If necessary, please leave your name and telephone number and someone from the study team will call you back as soon as possible.

**Please enter the date you completed the survey:** _ _/_ _/_ _ _ _
m m/ d d/ y y y y

## Background Information

To begin we would like to confirm your contact information.

| CONTACT INFORMATION | CORRECTIONS / UPDATES?  (Please fill in below) |
|---|---|
| | Respondent name: _____ |
| | Organization name: _____ |
| | Name of mentoring program: _____ |
| | Mailing address: _____ |
| | Telephone number: _____ |
| | E-mail address: _____ |

The following few questions are about your organization's budget, specifically both your total operating budget and your budget for *all* mentoring programs at your organization, and your budget *for your school-based mentoring programs*.  If you are a program operator and do not have this information, please obtain it from your organization's CFO or executive director.  The rest of the questions in the survey are about program operations and should be answered by the lead school-based mentoring program staff.

1.      What is the entire annual operating budget …

For your school-based mentoring program?                $_____

For all mentoring programs at your organization?        $_____

For your entire organization?                           $_____

2.      If your school-based mentoring efforts are funded by organizations other than the U.S. Department of Education, what kinds of organizations are they?  *(Check all that apply.)*

❑₁  Other government agencies
❑₂  Foundations
❑₃  Corporations
❑₄  Individual donors
❑₅  Other: _____

3.      Does your organization run any other programs for at-risk youth?

❑₁  Yes  →  *Go to question 4.*
❑₂  No  →  *Go to question 5.*

4.      If your organization runs other activities/programs for at-risk youth that are separate and distinct from the school-based mentoring program, what are they? *(Check all that apply.)*

    ❑₁  After-school programs (that are distinct from school-based mentoring program efforts that occur after school)        57/

    ❑₂  Tutoring        58/

    ❑₃  Counseling        59/

    ❑₄  Organization refers at-risk youth to other social service agencies        60/

    ❑₅  Organization provides support services (counseling, referrals, etc.) to the families of at-risk youth        61/

    ❑₆  Other mentoring program(s)        62/

    ❑₇  Other: _____        63/

                64-78/

5.      Describe your organization.  Is it…   *(Check one response.)*

    ❑₁  A chapter of Big Brothers Big Sisters, Inc.?        79/

    ❑₂  A school or school district?

    ❑₃  A non-profit/community based organization other than Big Brothers Big Sisters, Inc.?

    ❑₄  A faith-based organization?

    ❑₅  Other? _____

                80-94/

6.      Please describe any of the organizations with which you are partnering to run your school-based mentoring program.  Are you partnering with … *(Check all that apply.)*

    ❑₁  Our organization is not partnering with other organizations.        95/

    ❑₂  Faith-based organization(s)?        96/

    ❑₃  Big Brothers Big Sisters, Inc.?        97/

    ❑₄  Non-profit/community based organization(s) other than Big Brothers Big Sisters, Inc.?        98/

    ❑₅  The criminal justice system?        99/

    ❑₆  Local company(ies)?        100/

    ❑₇  Local college(s)/university(ies)?        101/

    ❑₈  Other? _____        102/

                103-117/

7.      How many paid staff people work on your organization's school-based mentoring program?

    7a. _____ Number of full-time staff, fully dedicated to your school-based mentoring program.        118-120/

    7b. _____ Number of part-time staff, or staff who are full-time, but who dedicate less than full-time to your school-based mentoring efforts.        121-123/

8.    Does your program employ a program coordinator specifically for your school-based mentoring program?

❑₁  Yes  → *Go to question 9.*                                                                  124/
❑₂  No   → *Go to question 11.*


9.    How many employees serve in this program coordinator role?  Please provide an answer in terms of full-time equivalents (FTEs) (i.e., 1 full-time person = 1 FTE; 1 full-time person and 1 half-time person = 1.5 FTEs; etc.).

_____  Number of FTEs                                                                         125-127/


**10.**   Please describe the relevant training and experience that the program coordinator has completed. (Check all that apply. If there is more than one person in this role, please answer for the most senior staff person in this position.)

**Education**

❑₁  High school degree or GED          128/
❑₂  Vocational degree or certification  129/
❑₃  2-year college degree              130/
❑₄  4-year college degree              131/
❑₅  Advanced (master's or higher) degree in education      132/
❑₆  Advanced degree in social work     133/

**Other experience**

❑₇  Prior experience teaching          134/
❑₈  Prior experience as a social worker   135/
❑₉  Prior experience working with volunteers      136/
❑₁₀ Prior experience working at a community-based organization    137-138/
❑₁₁ Prior experience working at a faith-based organization        139-140/


## Organization's Experience in School-based Mentoring

11.   For how many years has your organization been running a school-based mentoring program?

_____  Number of years of experience                                                         141-143/


12.   Which of the following best describes your school-based mentoring efforts? *(Check one response.)*

❑₁  School-based mentoring is the primary focus of our organization.                             144/
❑₂  School-based mentoring is important to our organization, but not our primary focus.
❑₃  School-based mentoring is one of many services that we offer youth.
❑₄  Other: _____
                                                                                                145-159/

13.   How many mentor/mentee matches were made in your school-based mentoring program in the past school year? *(Here, we define a "match being made" as a mentor and mentee having been paired. Note: if a mentor has more than two mentees this counts as two matches. The number of matches made may not equal the number of mentees served.)*

        _____  Number of matches                                                        160-162/

14.   Consider the length of mentor/mentee relationships in your school-based mentoring program.

      *(For the following question, please provide a number for each category. When you determine the length of a mentoring relationship, consider that relationship to have begun once mentees and mentors were matched, even if they never met in person. Please consider a relationship to have ended on the last day that the match met. For example, if Jane Mentor and John Mentee were matched on 9/1/06 and their last meeting was 3/1/07, their relationship would have lasted 6 months and you would indicate 6 to 9 months below. If matches lasted exactly 6 months indicate 6 to 9 months. If matches lasted 9 months indicate 9 to 12 months.)*

      14a.  By June 30th this year, what percentage of matches will have lasted each of the following time periods:

            ❑₁  Our program does not collect these data → *Go to question 14b.*          163/

| Percentage of matches | Duration | |
|---|---|---|
| _____ | Less than 3 months | 164-166/ |
| _____ | 3 to 6 months | 167-169/ |
| _____ | 6 to 9 months | 170-172/ |
| _____ | 9 to 12 months | 173-175/ |
| _____ | More than 12 months | 176-178/ |
| **100%** | **Total** | |

      14b.  By which month were more than 50% of mentees in your school-based mentoring program matched with a mentor? *(E.g., if 50% of mentees and mentors were matched in September and another 20% were matched in October, the answer would be "October".)*

            ❑₁  September            ❑₇  March             179-180/
            ❑₂  October             ❑₈  April
            ❑₃  November            ❑₉  May
            ❑₄  December            ❑₁₀ June
            ❑₅  January             ❑₁₁ July
            ❑₆  February            ❑₁₂ August

## Organization's Goals for Mentoring

15. To what extent is your school-based mentoring program focused on addressing each of the following goals? *(Check one response per row.)*

| | Our program is not focused on this at all. | Our program is a little focused on this. | Our program is moderately focused on this. | Our program is extremely focused on this. | |
|---|:---:|:---:|:---:|:---:|---|
| Improving mentees' self-esteem | ❑₁ | ❑₂ | ❑₃ | ❑₄ | 181/ |
| Providing mentees with general guidance | ❑₁ | ❑₂ | ❑₃ | ❑₄ | 182/ |
| Improving mentees' relationships with their parents/other caregivers | ❑₁ | ❑₂ | ❑₃ | ❑₄ | 183/ |
| Improving mentees' relationships with other adults in authority (teachers, principals, probation officers, etc.) | ❑₁ | ❑₂ | ❑₃ | ❑₄ | 184/ |
| Improving mentees' relationships with peers | ❑₁ | ❑₂ | ❑₃ | ❑₄ | 185/ |
| Improving mentees' attitudes towards school | ❑₁ | ❑₂ | ❑₃ | ❑₄ | 186/ |
| Improving mentees' academic performance in school | ❑₁ | ❑₂ | ❑₃ | ❑₄ | 187/ |
| Improving mentees' attendance | ❑₁ | ❑₂ | ❑₃ | ❑₄ | 188/ |
| Improving the likelihood that mentees will not drop out of school before graduating from high school | ❑₁ | ❑₂ | ❑₃ | ❑₄ | 189/ |
| Improving mentees' ability to plan for the future (to think about graduating from school, going to college, planning for jobs, etc.) | ❑₁ | ❑₂ | ❑₃ | ❑₄ | 190/ |
| Increasing mentees' likelihood of being engaged in their communities (participating in community service activities, etc.) | ❑₁ | ❑₂ | ❑₃ | ❑₄ | 191/ |
| Increasing mentees' ability to refrain from getting involved in gangs | ❑₁ | ❑₂ | ❑₃ | ❑₄ | 192/ |
| Increasing mentees' ability to refrain from engaging in violent activities | ❑₁ | ❑₂ | ❑₃ | ❑₄ | 193/ |
| Increasing mentees' ability to refrain from engaging in criminal activities | ❑₁ | ❑₂ | ❑₃ | ❑₄ | 194/ |
| Increasing mentees' ability to refrain from using drugs/alcohol | ❑₁ | ❑₂ | ❑₃ | ❑₄ | 195/ |

| | Our program is not focused on this at all. | Our program is a little focused on this. | Our program is moderately focused on this. | Our program is extremely focused on this. | |
|---|---|---|---|---|---|
| Increasing mentees' ability to refrain from high-risk sexual behaviors | ❏$_1$ | ❏$_2$ | ❏$_3$ | ❏$_4$ | 196/ |

16. Would it be accurate to say that, in your school-based mentoring program… *(Check one response.)*

     ❏$_1$ A. Mentors and mentees are encouraged to spend time together, but the organization does not encourage mentors and mentees to focus on particular issues? *(Go to question 19.)*    197/

     **OR**

     ❏$_2$ B. Program staff encourage mentors and mentees to engage in particular activities and to work on specific behaviors, but there is not a formal written curriculum for mentors? *(Please answer questions 17 and 18.)*

     **OR**

     ❏$_3$ C. Program staff encourage mentors and mentees to engage in particular activities and to work on specific behaviors, and there is a formal written curriculum for mentors (developed either commercially or by the program)? *(Please answer questions 17 and 18.)*

17. If the answer to question 16 is B or C: To what extent do your mentors and mentees in your school-based mentoring program focus on each of the following? *(Check one response per row.)*

| | Mentors and mentees are not focused on this at all. | Mentors and mentees are a little focused on this. | Mentors and mentees are moderately focused on this. | Mentors and mentees are extremely focused on this. | |
|---|---|---|---|---|---|
| Academics (working with mentees on homework, basic skills, etc.) | ❏$_1$ | ❏$_2$ | ❏$_3$ | ❏$_4$ | 198/ |
| Career exploration (educating mentees about the world of work, discussing career opportunities that mentees may wish to pursue, etc.) | ❏$_1$ | ❏$_2$ | ❏$_3$ | ❏$_4$ | 199/ |
| Goals that mentees establish for themselves | ❏$_1$ | ❏$_2$ | ❏$_3$ | ❏$_4$ | 200/ |
| Preventing mentee involvement with drugs/drug abuse | ❏$_1$ | ❏$_2$ | ❏$_3$ | ❏$_4$ | 201/ |
| Preventing early sexual activity | ❏$_1$ | ❏$_2$ | ❏$_3$ | ❏$_4$ | 202/ |
| Preventing truancy/ dropping out of school | ❏$_1$ | ❏$_2$ | ❏$_3$ | ❏$_4$ | 203/ |
| Preventing violence | ❏$_1$ | ❏$_2$ | ❏$_3$ | ❏$_4$ | 204/ |

| | Mentors and mentees are not focused on this at all. | Mentors and mentees are a little focused on this. | Mentors and mentees are moderately focused on this. | Mentors and mentees are extremely focused on this. | |
|---|---|---|---|---|---|
| Preventing gang involvement | $\Box_1$ | $\Box_2$ | $\Box_3$ | $\Box_4$ | 205/ |
| Other: _____ _____ | $\Box_1$ | $\Box_2$ | $\Box_3$ | $\Box_4$ | 206/ |

207-221/

18.    Which of these behaviors is the most important focus of your school-based mentoring program? *(Check one response.)*

$\Box_1$  Academics (working with mentees on homework, basic skills, etc.)                          222/

$\Box_2$  Career exploration (educating mentees about the world of work, discussing career opportunities that mentees may wish to pursue, etc.)

$\Box_3$  Goals that mentees establish for themselves

$\Box_4$  Preventing mentee involvement with drugs/drug abuse

$\Box_5$  Preventing early sexual activity

$\Box_6$  Preventing truancy/ dropping out of school

$\Box_7$  Preventing violence

$\Box_8$  Preventing gang involvement

$\Box_9$  Other: _____

223-237/

## Recruiting Mentors

19.    How does your school-based mentoring program recruit mentors? *(Check all that apply.)*

$\Box_1$  1. Through ads in local newspapers                                                  238/

$\Box_2$  2. Through public service announcements on local radio or TV stations              239/

$\Box_3$  3. By distributing flyers and brochures around the community                       240/

$\Box_4$  4. Through partnerships with local businesses                                      241/

$\Box_5$  5. Through partnerships with local faith-based organizations                       242/

$\Box_6$  6. Through partnerships with local universities/colleges                           243/

$\Box_7$  7. Through referrals from intermediary organizations (such as mentoring organization, volunteer centers or statewide recruitment campaigns)      244/

$\Box_8$  8. Through referrals from Board members                                            245/

$\Box_9$  9. Through referrals from mentors already involved in the program                  246/

$\Box_{10}$  10. Other: _____      247-248/

249-263/

20. Which of these strategies (1–10) is most effective for recruiting mentors? *(Select one strategy.)*

    \_\_\_\_\_ Strategy #

## Recruiting Mentees

21. What is the primary source of referrals for mentees in your school-based mentoring program? *(Check one response.*)

    ❏₁ Classroom teachers
    ❏₂ School guidance counselors
    ❏₃ Community-based organizations
    ❏₄ Faith-based organizations
    ❏₅ Social service agencies
    ❏₆ The criminal justice system
    ❏₇ Other *(Please list.)*_____

22. How many school districts have students participating in the school-based mentoring program in the past school year?

    \_\_\_\_\_ Number of districts

23. How many schools have students participating in your school-based mentoring program in the past school year?

    \_\_\_\_\_ Number of schools

## School-based Mentoring Program Mentee Demographics and Risk Factors

24. What is the total number of mentees participating in your school-based mentoring program that your organization runs?

    \_\_\_\_\_ Number of mentees

25. What percentage of mentees in the school-based mentoring program that your organization runs were in the following grades in the past school year?

☐₁ Our program does not collect these data → *Go to question 26.* <span style="float:right">291/</span>

| Percentage of mentees | Grade | |
|---|---|---|
| _____% | Grades K–3 | 292-294/ |
| _____% | Grade 4 | 295-297/ |
| _____% | Grade 5 | 298-300/ |
| _____% | Grade 6 | 301-303/ |
| _____% | Grade 7 | 304-306/ |
| _____% | Grade 8 | 307-309/ |
| _____% | Grades 9 and higher | 310-312/ |
| **100%** | **Total** | |

26. What's the racial/ethnic composition of mentees in the school-based mentoring program that your organization runs?

☐₁ Our program does not collect these data → *Go to question 27.* <span style="float:right">313/</span>

| Percentage of mentees | Race/Ethnicity | |
|---|---|---|
| _____% | American Indian or Alaska Native | 314-316/ |
| _____% | Asian | 317-319/ |
| _____% | Black or African American | 320-322/ |
| _____% | Hispanic or Latino | 323-325/ |
| _____% | Native Hawaiian or other Pacific Islander | 326-328/ |
| _____% | White | 329-331/ |
| **100%** | **Total** | |

27. What percent of mentees in your school-based mentoring program that your organization runs are not native English speakers?

☐₁ Our program does not collect these data → *Go to question 28.* <span style="float:right">332/</span>

_____% of mentees not native English speakers

28.  What percent of mentees in your school-based mentoring program that your organization runs is female?

_____ % female mentees

29.  What are the most common risk factors experienced by mentees in your school-based mentoring program? *(Check all that apply and rank (1–3) the top three, with 1 being the biggest risk factor.)*

| *Example:* |
| --- |
| ☑    _3_   A |
| ☑    _1_   B |
| ☑    ____   C |
| ☑    _2_   D |
| ☑    ____   E, etc. |

**Does it apply?**
**Check if yes.**  **Rank top 3**

| Check | code | Rank | code | Factor |
| --- | --- | --- | --- | --- |
| ☐1 | 339/ | ____ | 340/ | A. Parents abuse drugs/alcohol. |
| ☐2 | 341/ | ____ | 342/ | B. Child uses drugs/alcohol. |
| ☐3 | 343/ | ____ | 344/ | C. Child lives in extreme poverty. |
| ☐4 | 345/ | ____ | 346/ | D. Child has been neglected/abused. |
| ☐5 | 347/ | ____ | 348/ | E. Child has a learning disability/is developmentally delayed. |
| ☐6 | 349/ | ____ | 350/ | F. Child gets into frequent fights with peers. |
| ☐7 | 351/ | ____ | 352/ | G. Child has few/no positive adult role models. |
| ☐8 | 353/ | ____ | 354/ | H. Child has self-esteem problems. |
| ☐9 | 355/ | ____ | 356/ | I. Child is failing in school. |
| ☐10 | 357-358/ | ____ | 359/ | J. Child is in a gang. |
| ☐11 | 360-361/ | ____ | 362/ | K. Child has other behavioral problems. |
| ☐12 | 363-364/ | ____ | 365/ | L. Other: _____ |
|  |  |  |  | 366-380/ |

30. In your school-based mentoring program, what kinds of screening must **volunteers** go through before they are matched with mentees? *(Check all that apply.)*

    ❑₁ Written application process                                                          381/
    ❑₂ Personal interview                                                                   382/
    ❑₃ Background check (general)                                                           383/
    ❑₄ Background check—criminal records check                                             384/
    ❑₅ Background check—child and domestic abuse record checks                             385/
    ❑₆ Reference check                                                                     386/
    ❑₇ Other: _____                                          387/
                                                                                 388-402/

31. Are **mentors** in your school-based mentoring program required to participate in pre-match training/orientation?

    ❑₁ Yes  →  *Go to question 32.*                                                        403/
    ❑₂ No   →  *Go to question 34.*

32. If yes, what does that pre-match training/orientation for mentors include? *(Check all that apply.)*

    ❑₁ Opportunity to meet with mentees interested in having a mentor                       404/
    ❑₂ Introduction to the program (discussion of requirements for participation
       and program logistics)                                                              405/
    ❑₃ Cross-cultural sensitivity training                                                 406/
    ❑₄ Training in how to identify and address situations in which the mentee
       has been neglected or abused                                                        407/
    ❑₅ Training in encouraging mentees to plan for the future and to set long-term goals    408/
    ❑₆ Training in working with mentees on academic achievement                            409/
    ❑₇ Training in working with mentees on refraining from using drugs                      410/
    ❑₈ Training in working with mentees on refraining from engaging in
       other criminal behaviors                                                            411/
    ❑₉ Training in working with mentees on career preparation                              412/
    ❑₁₀ Training in working with mentees on drop-out reduction                            413-414/
    ❑₁₁ Training in working with mentees on refraining from engaging in violence          415-416/
    ❑₁₂ Training in working with mentees on refraining from gang involvement              417-418/
    ❑₁₃ Training in working with mentees on avoiding high risk sexual behaviors           419-420/
    ❑₁₄ Other: _____                                        421-422/
                                                                                 423-437/

33.   In your school-based mentoring program how many hours of pre-match training/orientation must each mentor receive before he or she is able to meet with his or her mentee?

☐₁  There is not a specific number of hours required → *Go to question 34.*       438/

\_\_\_\_\_  Number of hours       439-441/

34.   In your school-based mentoring program how many weeks, on average, does it take to get a volunteer screened and trained and matched with a mentee?

\_\_\_\_\_  Number of weeks       442-444/

35.   Are **mentees** in your school-based mentoring program provided with any sort of pre-match orientation/training?

☐₁  Yes  →  *Go to question 36.*       445/
☐₂  No  →  *Go to question 37.*

36.   If yes, what kind of pre-match orientation/training is provided to **mentees**?  *(Check all that apply.)*

☐₁  Introduction to the program (discussion of requirements for participation and program logistics)       446/
☐₂  Training in setting boundaries/resisting inappropriate advances from adults       447/
☐₃  Training in and help with setting goals for participation in the school-based mentoring program       448/
☐₄  Other: _____       449/
         450-464/

37.   How does your organization match mentees and mentors for your school-based mentoring program?  *(Check all that apply.)*

☐₁  Match mentees with mentors as soon as a mentor becomes available       465/
☐₂  Match "highest risk" mentees first       466/
☐₃  Aim to make same race matches       467/
☐₄  Aim to make same gender matches       468/
☐₅  Make matches based on personality-based assessment of what would constitute a good fit (survey of interests, etc.)       469/
☐₆  Mentees meet a pool of eligible mentors and can choose       470/
☐₇  Mentors meet a pool of eligible mentees and can choose       471/
☐₈  Other: _____       472/
         473-487/

## Parental Involvement

38. How does your program work with the parents/guardians of mentees in the school-based mentoring program? *(Check all that apply.)*

    ❑₁ Our program does not work with parents.                                              488/

    ❑₂ Parents/guardians meet with potential mentors before matches are made.               489/

    ❑₃ Parents/guardians play an active role in selecting particular mentors for
       their children.                                                                      490/

    ❑₄ Parents/guardians meet regularly with their children's mentors.                      491/

    ❑₅ Parents/guardians participate in group-activities with mentors, other
       youth in the program, and other parents.                                             492/

    ❑₆ Other: _____                                              493/

                                                                                    494-508/


## Participating Mentors

39. Who are the mentors in the school-based mentoring program? That is, what types of people have been recruited? *(Check all that apply.)*

    ❑₁ Teachers                                                                             509/

    ❑₂ Clergy                                                                               510/

    ❑₃ Employers of specific businesses or agencies                                         511/

    ❑₄ Retirees                                                                             512/

    ❑₅ General adult community members                                                      513/

    ❑₆ College students                                                                     514/

    ❑₇ High-school students                                                                 515/

    ❑₈ Other: _____                                              516/

                                                                                    517-531/


40. Are most of the mentors in your school-based mentoring program from one of these groups?

    ❑₁ Yes  →  *Go to question 41.*                                                         532/

    ❑₂ No   →  *Go to question 42.*


41. If most of the mentors are from one group, which group is it? *(Check one response.)*

    ❑₁ Teachers                                                                             533/

    ❑₂ Clergy

    ❑₃ Employers of specific businesses or agencies

    ❑₄ Retirees

    ❑₅ General adult community members

    ❑₆ College students

    ❑₇ High school students

    ❑₈ Other: _____

**Mentee/Mentor Activities**

42. During the school year, what is the minimum commitment required for mentees and mentors in your school-based mentoring program? *(Please answer all three of the following unless there is no specific minimum contact required.)*

    ☐₁ There is no specific minimum contact required. → *Go to question 43.*  549/

    42a. _____ Number of contacts per month  550-552/

    42b. _____ Number of hours per contact  553-555/

    42c. _____ Number of months  556-558/

43. What is the most common ratio for mentees and mentors in your school-based mentoring programs? *(Check one response.)*

    ☐₁ One mentor to one mentee  559/
    ☐₂ One mentor to several mentees
    ☐₃ One mentee to several mentors
    ☐₄ Several mentors to several mentees

44. Where do mentors and mentees in your school-based mentoring program meet most often? *(Check one response.)*

    ☐₁ At mentees' schools, during the school day  560/
    ☐₂ At the mentees' schools, after the school day is over
    ☐₃ At a faith-based organization
    ☐₄ At a community-based organization
    ☐₅ At a local company
    ☐₆ In the community (location selected by mentee and/or mentor)
    ☐₇ Somewhere else: _____
    561-575/

45. In what kinds of activities do most mentors and mentees in your school-based mentoring program engage? *(Check all that apply.)*

    ☐₁ Mentors and mentees spend time talking and "hanging out" together.  576/
    ☐₂ Mentors and mentees work on mentees' homework.  577/
    ☐₃ Mentors and mentees work on mentees' academic skills.  578/
    ☐₄ Mentors and mentees engage in community service activities.  579/
    ☐₅ Mentors and mentees visit the mentors' workplaces.  580/
    ☐₆ Mentors and mentees participate in group-activities sponsored by your organization (trips to local museums, libraries, ballgames, colleges, etc.)  581/
    ☐₇ Mentors meet with mentees' families.  582/
    ☐₈ Other: _____  583/

46.    Does your school-based mentoring program require or permit contact over the summer
       (outside of the traditional school year) between mentors and mentees?

☐₁   Yes, summer contact is required → *Go to question 47.*            <span>599/</span>
☐₂   Yes, summer contact is permitted (but not required) → *Go to question 47.*
☐₃   No → *Go to question 48.*

47.    If yes, how do mentors and mentees stay in touch over the summer? *(Check all that apply.)*

☐₁   Mentees and mentors continue to meet at school.                   600/
☐₂   Mentees and mentors continue to meet in the community.            601/
☐₃   Mentees and mentors are encouraged to exchange e-mails and/or letters and/or to talk
       on the telephone.                                               602/
☐₄   Mentees and mentors are encouraged to participate in group-activities sponsored by the
       organization.                                                   603/
☐₅   We don't encourage any specific behavior.                         604/
☐₆   Other: _____                                   605/

48.    After mentees "graduate" from (or age out of) your school-based mentoring program, is it
       possible for their mentors to continue to mentor them, with support from your organization as
       they do that?

☐₁   Yes                                                               621/
☐₂   No

## On-going Support and Training

49.    What kinds of **ongoing support** for mentors in your school-based mentoring program does
       your organization provide?  *(Check all that apply.)*

☐₁   Mentor/mentee meetings are supervised by program staff.           622/
☐₂   Mentors have access to social workers who are involved in the program
       and who can answer questions/address concerns.                  623/
☐₃   The organization hosts get-togethers where mentors can meet
       and discuss strategies for working effectively with mentees.    624/
☐₄   The organization sponsors listservs, mentoring chat rooms, or
       other on-line forums for mentors to support each other.         625/
☐₅   Mentors are required to participate in on-going trainings on a
       variety of issues.                                              626/
☐₆   Mentors have the opportunity to participate in on-going trainings
       on a variety of issues.                                         627/
☐₇   Other: _____                                   628/

❏₈ None, our organization focuses on providing support before the match is made and/or after mentoring relationships end.                                                          644/

50.    Are mentors in your school-based mentoring program required to report to program staff about their interactions with their mentees?

   ❏₁  Yes  →  *Go to question 51.*                                                          645/
   ❏₂  No  →  *Go to question 52.*


51.    If yes, how do they report on their activities?  *(Check all that apply.)*

   ❏₁  Mentors are required to provide program staff with detailed reports of each of their meetings with mentees, describing activities, etc.                    646/
   ❏₂  Mentors are required to check-in with program staff after every meeting, but they are not required to provide detailed information.                    647/
   ❏₃  Mentors are required to check-in with program staff periodically but not after every mentoring session.                    648/
   ❏₄  Mentors keep formal logs of contacts and times they meet with their mentees.   649/
   ❏₅  Other: _____                    650/
                                                                                                        651-665/


52.    Does your organization provide any **training opportunities** for mentors in your school-based mentoring program over the course of the program year (*in addition* to training provided as part of orientation)?

   ❏₁  Yes  →  *Go to question 53.*                                                          666/
   ❏₂  No  →  *Go to question 54.*


53.    If yes, what kinds of post-orientation trainings are offered over the program year? *(Check all that apply.)*

   ❏₁  Cross-cultural sensitivity training                                                    667/
   ❏₂  Training in how to identify and address situations in which the mentee has been neglected or abused                    668/
   ❏₃  Training in working with mentees on academic achievement           669/
   ❏₄  Training in working with mentees on refraining from using drugs     670/
   ❏₅  Training in working with mentees on refraining from engaging in other criminal behaviors                    671/
   ❏₆  Training in working with mentees on career preparation                672/
   ❏₇  Training in working with mentees on drop-out reduction                673/
   ❏₈  Training in working with mentees on refraining from engaging in violence   674/
   ❏₉  Training in working with mentees on refraining from gang involvement.   675/
   ❏₁₀ Training in working with mentees on avoiding high risk sexual behaviors   676-677/
   ❏₁₁ Other ongoing training. *(Please describe:)* _____   678-679/
                                                                                                        680-694/

## Information About Program Challenges/Needs for Technical Assistance

54.    How difficult is it for you to implement each of the following aspects of mentoring program operation in your school-based mentoring program? *(Check one response per row.)*

|  | Very difficult | Somewhat difficult | Not very difficult | Not at all difficult |  |
|---|---|---|---|---|---|
| Recruiting mentors | ❑1 | ❑2 | ❑3 | ❑4 | 695/ |
| Retaining mentors | ❑1 | ❑2 | ❑3 | ❑4 | 696/ |
| Recruiting mentees | ❑1 | ❑2 | ❑3 | ❑4 | 697/ |
| Retaining mentees | ❑1 | ❑2 | ❑3 | ❑4 | 698/ |
| Screening mentors (including background checks) | ❑1 | ❑2 | ❑3 | ❑4 | 699/ |
| Training mentors before they are matched | ❑1 | ❑2 | ❑3 | ❑4 | 700/ |
| Providing ongoing, post-match support and post-match training for mentors | ❑1 | ❑2 | ❑3 | ❑4 | 701/ |
| Hiring and retaining quality staff | ❑1 | ❑2 | ❑3 | ❑4 | 702/ |
| Fundraising | ❑1 | ❑2 | ❑3 | ❑4 | 703/ |
| Documenting program outcomes | ❑1 | ❑2 | ❑3 | ❑4 | 704/ |
| Obtaining appropriate insurance/liability coverage | ❑1 | ❑2 | ❑3 | ❑4 | 705/ |
| Other challenges *(Specify:)* _____ _____ 707-721/ | ❑1 | ❑2 | ❑3 | ❑4 | 706/ |

Please return this survey, in the enclosed pre-paid envelope, to:

> **Christina Dyous**
> **Spring Grantee Survey**
> **c/o Abt Associates Inc.**
> **55 Wheeler St.**
> **Cambridge, MA 02138**

**THANK YOU FOR PARTICIPATING IN THIS SURVEY!**

# Mentor Survey

# General Questions

**Please return by _____**

# Instructions

Please complete all questions; each question includes directions for recording your answer.

*You are sometimes told to skip over some questions in the survey. When this happens, you will see an arrow with a note that tells you what question to answer like this:*

   ❑₁ Yes
   ❑₂ No → *Go to question E4*

*Please complete this survey in terms of your experiences in the past school year.*

## Information about you

To begin, we would like to confirm your contact information. This information will help us be able to contact you and to mail you your incentive payment.

| CONTACT INFORMATION | CORRECTIONS / UPDATES? (Please fill in below) | |
|---|---|---|
| | Name: _____ | 10-34/ 35-69/ |
| | E-mail address: _____ | 70-119/ |
| | Home telephone #: _____ | 120-129/ |
| | Work telephone #: _____ | 130-139/ 140-143/ |
| | Cell phone #: _____ | 144-153/ |
| | | 154-188/ 189-223/ 224-225/ |
| | Mailing Address: _____ | 226-230/ |

## A. Information About Your Involvement in the Student Mentoring Program

The next questions are about your involvement in the Student Mentoring Program.

A1.    How did you hear about the program?  Was it through … **(Check all that apply.)**

☐₁ Ads in local newspapers                                                        231/
☐₂ Public service announcements on local radio or TV stations                       232/
☐₃ Flyers/brochures distributed around the community                               233/
☐₄ My employer                                                                    234/
☐₅ My church/synagogue/mosque/ other faith-based organization                      235/
☐₆ My university/college                                                           236/
☐₇ An intermediary organization (such as a mentoring organization, a volunteer center or through a statewide recruitment campaign)                 237/
☐₈ Referrals from Board members at the organization that runs the program          238/
☐₉ Mentors already involved in the program                                        239/
☐₉₅ Other: _____                                        240-241
                                                              242-243/  244-245/  246-247/

A2. How important were the following in your decision to volunteer in the Student Mentoring Program?

| | Check one box for each row below … | | | |
|---|---|---|---|---|
| | Not Important | Not very important | Somewhat important | Very important |
| a. Volunteering makes me feel needed. | ☐₁ | ☐₂ | ☐₃ | ☐₄ 248/ |
| b. I feel compassion for people who need help. | ☐₁ | ☐₂ | ☐₃ | ☐₄ 249/ |
| c. People I respect consider volunteering an important activity. | ☐₁ | ☐₂ | ☐₃ | ☐₄ 250/ |
| d. Volunteering provides me with new perspective on things. | ☐₁ | ☐₂ | ☐₃ | ☐₄ 251/ |
| e. Volunteering is an important aspect of my faith. | ☐₁ | ☐₂ | ☐₃ | ☐₄ 252/ |
| f. I am eager to participate in an activity where I have the opportunity to meet new people who are also interested in volunteering. | ☐₁ | ☐₂ | ☐₃ | ☐₄ 253/ |
| g. I want to give back to my community. | ☐₁ | ☐₂ | ☐₃ | ☐₄ 254/ |
| h. I want to gain experience working with youth. | ☐₁ | ☐₂ | ☐₃ | ☐₄ 255/ |
| i. I want to gain experience working in/with schools. | ☐₁ | ☐₂ | ☐₃ | ☐₄ 256/ |
| j. I have always wanted to be a mentor. | ☐₁ | ☐₂ | ☐₃ | ☐₄ 257/ |
| k. Participating in the Student Mentoring Program fits well with my schedule. | ☐₁ | ☐₂ | ☐₃ | ☐₄ 258/ |
| l. I had heard about this program/sponsoring organization and wanted to volunteer there. | ☐₁ | ☐₂ | ☐₃ | ☐₄ 259/ |
| m. Other:_____ _____ 261-262/ 263-264/ 265-266/ | ☐₁ | ☐₂ | ☐₃ | ☐₄ 260/ |

A3. Have you volunteered as a mentor before participating in the Student Mentoring Program?      267/

☐₁ Yes
☐₂ No

A4. How much contact had you had, prior to your participation in the Student Mentoring Program, with students in grades 4–8?      268/

☐₁ None
☐₂ Very little
☐₃ Some
☐₄ A lot

A5. From the list below, please rank order (1-3, with 1 being the most important) the three most important things you hope to accomplish with your student through the Student Mentoring Program.

**Example:**

| | |
|---|---|
| ____3___ | Answer choice A |
| ____2___ | Answer choice B |
| _____ | Answer choice C |
| ____1___ | Answer choice D, etc. |

| # | Rank top 3 below … | |
|---|---|---|
| _____ | Increase the student's self-esteem | 269-270/ |
| _____ | Provide student with general guidance | 271-272/ |
| _____ | Improve the student's relationships with his/her parents or caregivers | 273-274/ |
| _____ | Improve the student's relationships with other adults in authority (teachers, principals, probation officers, etc.) | 275-276/ |
| _____ | Improve the student's relationships with peers | 277-278/ |
| _____ | Improve the student's attitudes towards school | 279-280/ |
| _____ | Improve the student's academic performance in school | 281-282/ |
| _____ | Improve student's attendance | 283-284/ |
| _____ | Improve the likelihood that the student will not drop out of school before graduating from high school | 285-286/ |
| _____ | Improve student's ability to plan for the future (to think about graduating from school, going to college, planning for jobs, etc.) | 287-288/ |
| _____ | Increase the likelihood that the student will be engaged in his/her community (participating in community service activities, etc.) | 289-290/ |
| _____ | Decrease the likelihood that the student will become involved in gangs | 291-292/ |
| _____ | Decrease the likelihood that the student will engage in violence | 293-294/ |
| _____ | Decrease the likelihood that the student will engage in criminal activities | 295-296/ |
| _____ | Decrease the likelihood that the student will use drugs/alcohol | 297-298/ |
| _____ | Decrease the likelihood that the student will engage in high-risk sexual behaviors | 299-300/ |
| _____ | Other: _____ <br> 303-304/ 305-306/ 307-308/ | 301-302/ |

## B.    Pre-Match Activities

B1.    What kind of screening did you receive before you were matched with your student?  **(Check all that apply.)**

❑₁  Written application process                                                                                                309/
❑₂  Personal interview                                                                                                              310/
❑₃  Background/criminal records/child and domestic abuse records check              311/
❑₄  Reference check
312/
❑₅  No screening was done prior to being matched  with my student               313/
❑₉₅  Other: _____               314-315/

316-317/   318-319/   320-321/

B2.    What kind of orientation and training did you receive before you were matched with your student?  **(Check all that apply.)**

❑₁  Opportunity to meet with students interested in having a mentor                     322/
❑₂  Introduction to the program (discussion of requirements for participation and program
      logistics)                                                                                                                            323/
❑₃  Cross-cultural sensitivity training                                                                             324/
❑₄  Training in encouraging students to plan for the future and to set long-term goals    325/
❑₅  Training in how to identify and address situations in which the student has been
      neglected or abused                                                                                                     326/
❑₆  Training in working with students on academic achievement                            327/
❑₇  Training in working with students on refraining from using drugs                      328/
❑₈  Training in working with students on refraining from engaging in other criminal
behaviors                                                                                                                           329/
❑₉  Training in working with students on career preparation                                  330/
❑₁₀ Training in working with students on drop-out reduction                              331-332/
❑₁₁ Training in working with students on refraining from engaging in violence        333-334/
❑₁₂ Training in working with students on refraining from gang involvement           335-336/
❑₁₃ Training in working with students on avoiding high risk sexual behaviors        337-338/
❑₁₄ No orientation or training was provided prior to being matched  with
      my student → *Go to question B4*                                                                          339-340/
❑₉₅  Other: _____               341-342/

343-344/   345-346/   347-348/

B3.    About how many hours of training or orientation did you receive before you met with your student for the first time?


_____ Number of hours                                                                                          349-352/

B4.     How many weeks did it take you to get screened and trained before you met with your student for the first time?

_____ Number of weeks                                                                           353-356/

## C.     Program Participation

C1.     Are you currently mentoring a student in this program?

❑₁  Yes → *Go to question C3*
        357/
❑₂  No, but I met with a student in this program
❑₃  No, I never met with any student in this program → *Go to question C7*

C2.     If you are not currently mentoring a student but did previously meet with a student in this program, how long ago did you stop mentoring? (When was the last meeting with your last student?)

❑₁  Last month (within the last four weeks)
        358/
❑₂  One to less than three months ago
❑₃  Three to less than six months ago
❑₄  Six or more months ago

C3.     Describe the nature of most of your mentoring sessions.  **(Check one answer.)**

❑₁  I shared one student with another mentor
        359/
❑₂  I mentored two or more students at a time with one or more other mentors
❑₃  I mentored one student at a time
❑₄  I mentored two or more students at a time

C4.     Since you became active in the program, what kinds of ongoing support has the organization that runs the mentoring program provided you and your student(s)?  **(Check all that apply.)**

❑₁  Mentor/student meetings were supervised by program staff
        360/
❑₂  Mentors had access to social workers or program staff who were involved in the program and          who could answer questions/address concerns
        361/
❑₃  The organization hosted get-togethers where mentors could meet one another and discuss strategies for working effectively with students
        362/

- ☐<sub>4</sub> The organization sponsored listservs, mentoring chat rooms, or other on-line forums for mentors to support each other
  363/
- ☐<sub>5</sub> Other:
  _____364/

365-366/ 367-368/ 369-370/

- ☐<sub>6</sub> None, the host organization focused instead on providing support before the match was made and/or after mentoring relationships ended
  371/

372/Blank

C5. About how often did you see or talk with a mentoring program supervisor about how things were going?

- ☐<sub>1</sub> More than once a week
  373/
- ☐<sub>2</sub> Once a week
- ☐<sub>3</sub> Once every two or three weeks
- ☐<sub>4</sub> Once a month
- ☐<sub>5</sub> Less than once a month
- ☐<sub>6</sub> Never

C6. Was this contact…

- ☐<sub>1</sub> Required?
  374/
- ☐<sub>2</sub> Strongly encouraged but not required?
- ☐<sub>3</sub> Not required and not strongly encouraged?
- ☐<sub>4</sub> Discouraged?

375/Blank

C7. Did the organization provide any training opportunities for mentors to participate in over the course of the program year (in addition to the training provided as part of orientation)?

- ☐<sub>1</sub> Yes
  376/
- ☐<sub>2</sub> No → *Go to question D1*

C8. If yes, what kinds of trainings were offered over the program year, not including the training provided as part of orientation? **(Check all that apply.)**

- ☐<sub>1</sub> Cross-cultural sensitivity training
  377/
- ☐<sub>2</sub> Training in how to identify and address situations in which the student has been neglected or abused
  378/
- ☐<sub>3</sub> Training in encouraging students to plan for the future and to set long-term goals
  379/
- ☐<sub>4</sub> Training in working with student on homework
  380/

❏₅ Training in working with student on academic skills

❏₆ Training in working with students on academic achievement

❏₇ Training in working with students on refraining from using drugs

❏₈ Training in working with students on refraining from engaging in other criminal behaviors.

❏₉ Training in working with students on career preparation

❏₁₀ Training in working with students on drop-out reduction

❏₁₁ Training in working with students on refraining from engaging in violence

❏₁₂ Training in working with students on refraining from gang involvement.

❏₁₃ Training in working with students on avoiding high-risk sexual behaviors.

❏₉₅ The organization provides other ongoing training. *(Please describe:)*

_____ _____

_____

_____

## D. Your Perceptions of the Student Mentoring Program

D1.    Please rate the quality of each of the following components of the Student Mentoring Program.

| | Check one box for each row below … | | | | | |
|---|---|---|---|---|---|---|
| | Extremely Poor | Poor | Fair | Good | Excellent | Not Applicable |
| Mentor screening process | ❏₁ | ❏₂ | ❏₃ | ❏₄ | ❏₅ | ❏₆ 402/ |
| Pre-match training/orientation | ❏₁ | ❏₂ | ❏₃ | ❏₄ | ❏₅ | ❏₆ 403/ |
| Process of matching students and mentors | ❏₁ | ❏₂ | ❏₃ | ❏₄ | ❏₅ | ❏₆ 404/ |
| Ongoing support from agency staff | ❏₁ | ❏₂ | ❏₃ | ❏₄ | ❏₅ | ❏₆ 405/ |
| Special group events for mentors and students | ❏₁ | ❏₂ | ❏₃ | ❏₄ | ❏₅ | ❏₆ 406/ |

| | Extremely Poor | Poor | Fair | Good | Excellent | Not Applicable |
|---|---|---|---|---|---|---|
| Mentoring program curriculum (A mentoring program curriculum is a particular focus on either discouraging or encouraging certain specified behaviors or on improving academic performance.) | ❑1 | ❑2 | ❑3 | ❑4 | ❑5 | ❑6  407/ |
| Ongoing support from school staff | ❑1 | ❑2 | ❑3 | ❑4 | ❑5 | ❑6  408/ |
| Appropriate support from staff when mentor/student relationship terminated | ❑1 | ❑2 | ❑3 | ❑4 | ❑5 | ❑6  409/ |

410-
419/Blank

D2.    Now, please assess how important each component was to your experience in the program.

| | Not Important | Somewhat Important | Very Important | Essential |
|---|---|---|---|---|
| Mentor screening process | ❑1 | ❑2 | ❑3 | ❑4  420/ |
| Pre-match training/orientation | ❑1 | ❑2 | ❑3 | ❑4  421/ |
| Process of matching students and mentors | ❑1 | ❑2 | ❑3 | ❑4  422/ |
| Ongoing support from agency staff | ❑1 | ❑2 | ❑3 | ❑4  423/ |
| Special group events for mentors and students | ❑1 | ❑2 | ❑3 | ❑4  424/ |
| Mentoring program curriculum | ❑1 | ❑2 | ❑3 | ❑4  425/ |
| Appropriate support from staff when mentor/student relationship terminated | ❑1 | ❑2 | ❑3 | ❑4  426/ |
| Ongoing support from school staff | ❑1 | ❑2 | ❑3 | ❑4  427/ |

Check one box for each row below …

D3.    How satisfied are you with your experience in the Student Mentoring Program?

❑1 Very satisfied
428/
❑2 Somewhat satisfied
❑3 Neutral
❑4 Somewhat dissatisfied

❑$_5$ Very dissatisfied

D4. What could have made your experience in this program better?  **(For this question, we are interested in learning about: 1) all of the things that could have made your experience in the program better <u>and</u> 2) about the top 3 things that could have improved your experience.  To answer this question, please first check all that apply and then rank (1-3) the top 3 things that could have enhanced your experience, with 1 being the most important.)**

**Example:**

| | | |
|---|---|---|
| ☑$_1$ | _3___ | Answer choice A |
| ☑$_2$ | _2___ | Answer choice B |
| ☑$_3$ | ____ | Answer choice C |
| ☑$_4$ | _1___ | Answer choice D, etc. |

| **Does it apply?** Check if yes | **Rank top 3** | |
|---|---|---|
| ❑$_1$ 429/ | ____ 430-431/ | More/better training before being matched with your student |
| ❑$_2$ 432/ | ____ 433-434/ | More frequent meetings with student |
| ❑$_3$ 435/ | ____ 436-437/ | More frequent contact with program staff |
| ❑$_4$ 438/ | ____ 439-440/ | Less high-risk students in the program |
| ❑$_5$ 441/ | ____ 442-443/ | More supports for my student outside of the program.  (My student needs social services that s/he is not getting.) |
| ❑$_6$ 444/ | ____ 445-446/ | More/better group activities |
| ❑$_7$ 447/ | ____ 448-449 | More opportunities to meet with other mentors |
| ❑$_8$ 450/ | ____ 451-452/ | More opportunities to meet with student's teachers |
| ❑$_9$ 453/ | ____ 454-455/ | More opportunities to meet with student's parents |
| ❑$_{10}$ 456-457/ | ____ 458-459/ | More support or supervision for mentors |
| ❑$_{11}$ 460-461/ | ____ 462-463/ | More training in _____ _____ _____ 464-465/ 466-467/ 468-469/ |
| ❑$_{95}$ 470-471/ | ____ 472-473/ | Other *(Please describe:)* _____ _____ 474-475/ 476-477/ 478-479/ |

D5. What were the biggest challenges in participating in the Student Mentoring Program? **(For this question, we are interested in learning about: 1) all of the challenges you experienced <u>and</u> 2) about the top 3 challenges you faced.  To answer this question, please first check all that apply and then rank (1-3) the top 3 challenges, with 1 being the biggest challenge.)**

**Example:**

| | | |
|---|---|---|
| ☑₁ | __3___ | Answer choice A |
| ☑₂ | __2___ | Answer choice B |
| ☑₃ | _____ | Answer choice C |
| ☑₄ | __1___ | Answer choice D, etc. |

| **Does it apply? Check if yes** | **Rank top 3:** | |
|---|---|---|
| ☐₁ 480/ | _____ 481-482 | It was hard for me to make the time to meet regularly. |
| ☐₂ 483/ | _____ 484-485/ | Pre-match training was ineffective/insufficient. |
| ☐₃ 486/ | _____ 487-488/ | It took too long/was too labor intensive to go through pre-match screening and orientation. |
| ☐₄ 489/ | _____ 490-491/ | I needed more training than I received.  It would have been particularly helpful to have training in_____ _____.*(Please fill in the blank.)* <br> 492-493/ 494-495/ 496-497/ |
| ☐₅ 498/ | _____ 499-500/ | I didn't get enough support from program staff. |
| ☐₇ 501/ | _____ 502-503/ | My student(s) didn't get enough support from program staff. |
| ☐₈ 504/ | _____ 505-506/ | I didn't know what to do with my student(s).  We needed more structured activities. |
| ☐₉ 507/ | _____ 508-509/ | It was difficult to establish a relationship with my student(s). |
| ☐₁₀ 510-511/ | _____ 512-513/ | My student(s) often didn't show up for our meetings. |
| ☐₁₁ 514-515/ | _____ 516-517/ | My student(s) pressured me to get more involved in his or her life than I felt comfortable doing. |
| ☐₁₂ 518-519 | _____ 520-521/ | My student(s) and I were supposed to work on particular skills and behaviors.  Working on those things wasn't fun/interesting. |
| ☐₁₃ 522-523/ | _____ 524-525/ | My student(s) seemed embarrassed by the fact that he or she was in this program. |
| ☐₁₄ 526-527/ | _____ 528-529/ | My student(s) had problems that were too big for me to handle.  I felt overwhelmed. |

☐₁₅ <sub>530-531/</sub> _____ <sub>532-533/</sub> I really did not have any challenges participating in the program.

☐₉₅ <sub>534-535/</sub> _____ <sub>536-537/</sub> Other *(Please describe:)*

_____

_____ <sub>538-539/ 540-541/ 542-543</sub>

<sub>544-</sub>

<sub>553/Blank</sub>

## E. Your Plans for the Future

E1. Do you plan to participate in the Student Mentoring Program next year?

☐₁ Yes

<sub>554/</sub>

☐₂ No → *Go to question F1*

☐₉₇ Don't know

E2. Do you plan to participate next year with your current student?

☐₁ Yes

<sub>555/</sub>

☐₂ No

☐₉₇ Don't know

## F. Information About You

F1. Are you…

☐₁ Male

☐₂ Female <sub>556/</sub>

F2. How old are you? <sub>557-559/</sub>

_____ years

F3. Are you Hispanic or Latino? <sub>560/</sub>

☐₁ Yes, Hispanic or Latino

☐₂ No, not Hispanic or Latino

F4. What is your race? **(Please check one or more.)**

☐₁ American Indian or Alaskan Native <sub>561/</sub>

☐₂ Asian <sub>562/</sub>

☐₃ Black or African American <sub>563/</sub>

☐₄ Native Hawaiian or other Pacific Islander <span style="color:gray">564/</span>

☐₅ White <span style="color:gray">565/</span>

F5. Are you a native English speaker? <span style="color:gray">566/</span>

    ☐₁ Yes → *Go to question F6*

    ☐₂ No

       F5a.    If no, what is your first language?

           First Language: _____ <span style="color:gray">567-568/</span>

<span style="color:gray">569-570/</span>

F6. Are you married or living with a partner? <span style="color:gray">571/</span>

    ☐₁ Yes, I am married or living with a partner.

    ☐₂ No, I am not married or living with a partner.

F7. Do you have children? <span style="color:gray">572/</span>

    ☐₁ Yes

    ☐₂ No → *Go to question F8*

    F7a. If you have children:

| | **Fill # below …** |
|---|---|
| a. How many live with you all or most of the time? | # _____ <span style="color:gray">573-574/</span> |
| b. How many of those that live with you are ages 0 – 5? | # _____ <span style="color:gray">575-576/</span> |
| c. How many of those that live with you are ages 6 – 11? | # _____ <span style="color:gray">577-578/</span> |
| d. How many of those that live with you are ages 12 – 18? | # _____ <span style="color:gray">579-580/</span> |

F8. What is your employment status *(Please select one answer. If you are a student and employed, please check that you are a full-time student.)*

- ☐₁ I am employed full-time
- ☐₂ I am employed part-time
- ☐₃ I am a full-time student → *Go to question F9*
- ☐₄ I am retired → *Go to question F9*
- ☐₅ I am not employed outside of the home → *Go to question F9*

F8a. In what field do you work?

- ☐₁ Agriculture
  582/
- ☐₂ The Arts/creative arts/performing arts/writing/music/dance/photography/film/video
- ☐₃ Construction
  584/
- ☐₄ Business
  585/
- ☐₅ Computers/technology/science
  586/
- ☐₆ Education/teaching
  587/
- ☐₇ Environmental
  588/
- ☐₈ Healthcare/health-related
  589/
- ☐₉ Law
  590/
- ☐₁₀ Military
  591-592/
- ☐₁₁ Public Safety
  593-594/
- ☐₁₂ Social/community work
  595-596/
- ☐₉₅ Other_____
  597-598/

F9. How much formal education have you completed?

- ☐₁ Some high school
- ☐₂ High school graduate or a GED
- ☐₃ Vocational or technical school certification

❏₄  Some college

❏₅  2 year college degree

❏₆  4 year college degree

❏₇  Some post-graduate study

❏₈  Advanced degree

F10.  Please record the date you completed this questionnaire.

|___|___|     |___|___|      |___|___|___|___|
   Month          Day                Year
   607-608/         609-610/           611-614/

*If you have not completed your Student-Specific Questionnaire(s), please do so at this time.*

*If you have completed all of the questionnaire(s), please return the surveys, in the enclosed pre-paid envelope, to:*

---

**Mentor Survey**
**C/O Abt Associates Inc.**
**55 Wheeler St.**
**Cambridge, MA 02138**

---

**THANK YOU FOR PARTICIPATING IN THIS SURVEY!**

# Mentor Survey

## Student Specific Questions

**Please complete this survey for:**

(insert student name label)

**Please return by**

# Instructions

Please complete these questions based on your experience with the student whose name is on the front of this booklet.  Please complete a booklet of student specific questions for each student with whom you were paired in the Student Mentoring Program during the past school year.  Please complete all questions.  Each question includes directions for recording your answer.

*You are sometimes told to skip over some questions in the survey.  When this happens, you will see an arrow with a note that tells you what question to answer like this:*

      ❑₁  Yes
      ❑₂  No → *Go to question 4*

*If you have any questions about how to complete the survey, or if there is another student for whom we did not provide a survey, please call the following toll-free number:*
*1-866-534-9161.  If needed, please leave your name and telephone number, and someone from the study team will call you back as soon as possible.*

**Student-Specific Questions**

1.  Did you ever meet with this student?

    ❑₁ Yes → *Go to question 6*
    10/
    ❑₂ No

2.  If no, who initiated the end of the match before you and your student met?

    ❑₁ I did.
    11/
    ❑₂ The student did. → *Go to question 4*
    ❑₃ The agency or school did. → *Go to question 4*

3.  If **you** ended the match, why did you do so?  **(Check all that apply.)**

    ❑₁ It was too much of a time commitment
    12/
    ❑₂ My student had needs that I could not meet; I felt overwhelmed
    13/
    ❑₃ The program was disorganized/run poorly
    14/
    ❑₄ I did not feel welcome at my student's school
    15/
    ❑₉₅ Other *(Please describe:)* _____
    _____ 16-17/

    18-19/  20-21/
    22-23/

4.  How disappointed were **you** that the match ended?
    24//

    ❑₁ Not disappointed at all
    ❑₂ A little disappointed
    ❑₃ Fairly disappointed
    ❑₄ Extremely disappointed

5.  How disappointed did **your student** appear to be that the match ended?

    ❑₁ Not disappointed at all
    25/
    ❑₂ A little disappointed
    ❑₃ Fairly disappointed
    ❑₄ Extremely disappointed
    ❑₅ Couldn't tell/not sure

**If you never met with this student,** you have now completed this section of the survey. If you had a mentoring relationship with any other students in this program, whose name appears on one of the attached survey cover pages, please complete the separate questionnaire for that/those student(s) in terms of your experience with that youth. If you met with another student whose name is not on the attached survey cover pages, please do not complete the rest of the survey (this survey captures the experiences of any youth and mentors who are part of the evaluation of the Student Mentoring Program, and not necessarily every student with whom you may have met).

If you have filled out a questionnaire for each student with whom you have/had a mentoring relationship, the survey is now complete. Please place all completed questionnaires in the pre-paid envelope provided.


Thank you for participating!

6.	Please check the month that you and your student were **matched**.

*Consider you and your student to be matched once you knew his or her name and expected to meet with him or her.*

Month you **were matched** with your student

| | |
|---|---|
| _____₁ | September |
| _____₂ | October |
| _____₃ | November |
| _____₄ | December |
| _____₅ | January |
| _____₆ | February |
| _____₇ | March |
| _____₈ | April |
| _____₉ | May |
| _____₁₀ | June |

26-
27/

6a.	Please check the month that you **stopped** or you **expect to stop** mentoring your student.

*Consider the relationship to have ended on the last day that you met with that student. If you plan to mentor this student beyond June, please check June as the month you expect to stop mentoring your student.*

Month you **stopped/expect to stop** mentoring your student

| | |
|---|---|
| _____₁ | September |
| _____₂ | October |
| _____₃ | November |
| _____₄ | December |
| _____₅ | January |
| _____₆ | February |
| _____₇ | March |
| _____₈ | April |
| _____₉ | May |
| _____₁₀ | June |

28-
29/

7.	During the past school year, on average, how often per month did you have in-person contact with your student?

_____ Average number of in-person contacts per month

30-
33/

8.	On average, how long, in minutes, was each in-person meeting with your student?

_____ Number of minutes

34-
37/

9.	Was the frequency of in person meetings with your student consistent throughout your experience in the program?   **(Check one answer.)** <sub>38-</sub>
<sub>39/</sub>

❑₁ Yes, we met about the same number of times per month throughout the program.

❑₂ No, initially we met in person regularly; at the end of the school year, we were meeting less, although we were in regular contact even when we didn't meet in person. (We e-mailed or spoke on the phone.)

❑₃ No, initially we met in person regularly; at the end of the school year we were meeting less frequently and we did not talk on the phone or e-mail when we missed meetings.

❑₄ No, initially we met sporadically, but as our relationship developed, we met more and more   frequently.

❑₉₅ Other (*Specify*): _____

_____ <sub>40-41/   42-43/</sub>
<sub>44-45/</sub>

10.	During the past school year, on average, how often per month did you have **other (DO NOT INCLUDE IN-PERSON CONTACTS)** kinds of contact with your student, such as, telephone, e-mail, fax, etc?

_____ Average number of other (not in-person) contacts per month <sub>46-</sub>
<sub>49/</sub>

11.	Rate your relationship with your student.

❑₁ *Extremely positive*.  We had a terrific, trusting relationship and were very close. <sub>50/</sub>

❑₂ *Somewhat positive*.  We had a good relationship and were moderately close.

❑₃ *Fair*.  We got along OK, but were not very close.

❑₄ *Poor*.  We really didn't connect.

12.	Where did you and your student meet most often?

❑₁ At school, during the school day <sub>51</sub>

❑₂ At school, after the school day <sub>52/</sub>

❑₃ At a community based organization <sub>53/</sub>

❑₄ At a faith-based organization <sub>54/</sub>

❑₅ At a local company <sub>55/</sub>

❑₆ In the community (location selected by student and mentor) <sub>56/</sub>

❑₉₅ Somewhere else *(Please describe:)* _____ <sub>57-</sub>
<sub>58/</sub>

<sub>59-60/  61-62/  63-64/</sub>

13.	What activities did you and your student do in the Student Mentoring Program?  **(Check all that apply.)**

❑₁  We spent time talking and "hanging out" together
<sub>65/</sub>

❑₂  We worked on the student's homework/academic skills
<sub>66/</sub>

❑₃  We engaged in community service activities
<sub>67/</sub>

☐₄ We visited my workplace
68/

☐₅ We participated in group activities sponsored by the organization (trips to local museums, 69/
libraries, ball games, colleges, etc.)

☐₆ I met with my student's family
70/

☐₉₅ Other (*Specify*): _____
71-72/

73-74/ 75-76/
77-78/

14. What strategies did you use to support your student? **(For this question, we are interested in learning both what kinds of strategies mentors used to support their students (in general) and what strategies they used most often. To answer this question, please first check all options that apply, then rank (1-3) the 3 strategies you used most frequently, with 1 indicating the most frequent strategy)**

**Example:**

| | | |
|---|---|---|
| ☑₁ | _3___ | Answer choice A |
| ☑₂ | _2___ | Answer choice B |
| ☑₃ | ____ | Answer choice C |
| ☑₄ | _1___ | Answer choice D, etc. |

| Does it apply? Check if yes | Rank top 3 | |
|---|---|---|
| ☐₁ 79/ | ____ 80/ | Listened and was a friend to the student |
| ☐₂ 81/ | ____ 82/ | Provided the student with a consistent and supportive adult role-model |
| ☐₃ 83/ | ____ 84/ | Provided the student with constructive criticism about his or her behavior |
| ☐₄ 85/ | ____ 86/ | Praised the student |
| ☐₅ 87/ | ____ 88/ | Shared my experiences and discuss how they have affected my life |
| ☐₆ 89/ | ____ 90/ | Exposed the student to new things |
| ☐₇ 91/ | ____ 92 | Set goals or standards for the student or helped the student set goals for him or herself |
| ☐₉₅ 93-94/ | ____ 95/ | Other. *(Please describe:)*_____ 96-97/ 98-99/ 100-101/ |

15. When you met with your student, how often did you do each of the following? **(Check the appropriate box for each item.)**

| | Check one box for each row below … | | | |
|---|---|---|---|---|
| | **Never** | **Some-times** | **Most of the time** | **Almost always** |
| Engaged in casual conversation | ☐₁ | ☐₂ | ☐₃ | ☐₄ 102/ |
| Talked about student's personal problems | ☐₁ | ☐₂ | ☐₃ | ☐₄ 103/ |
| Talked about student's aspirations for the future (career plans, college plans, etc.) | ☐₁ | ☐₂ | ☐₃ | ☐₄ 104/ |
| Talked about student's relationships with parents | ☐₁ | ☐₂ | ☐₃ | ☐₄ 105/ |
| Talked about student's relationships with teachers/other adults in authority | ☐₁ | ☐₂ | ☐₃ | ☐₄ 106/ |
| Talked about student's relationships with peers | ☐₁ | ☐₂ | ☐₃ | ☐₄ 107/ |
| Worked on academic skills | ☐₁ | ☐₂ | ☐₃ | ☐₄ 108/ |
| Worked on homework | ☐₁ | ☐₂ | ☐₃ | ☐₄ 109/ |
| Engaged in community service | ☐₁ | ☐₂ | ☐₃ | ☐₄ 110/ |
| Talked about the importance of completing high school | ☐₁ | ☐₂ | ☐₃ | ☐₄ 111/ |
| Talked about the risks associated with alcohol/drug use | ☐₁ | ☐₂ | ☐₃ | ☐₄ 112/ |
| Talked about the risks associated with engaging in violence/criminal activities | ☐₁ | ☐₂ | ☐₃ | ☐₄ 113/ |
| Talked about the risks associated with gang involvement | ☐₁ | ☐₂ | ☐₃ | ☐₄ 114/ |
| Talked about the risks associated with high-risk sexual activity | ☐₁ | ☐₂ | ☐₃ | ☐₄ 115/ |

116-

117/Blank

16. When you met with your student, how often did each of the following happen? **(Check the appropriate box for each item.)**

| | Check one box for each row below … | | | |
|---|---|---|---|---|
| | **Never** | **Some-times** | **Most of the time** | **Almost always** |
| Your student seemed really glad to see you. | ☐₁ | ☐₂ | ☐₃ | ☐₄ 118/ |

| | Never | Some-times | Most of the time | Almost always |
|---|---|---|---|---|
| Your student confided in you. | $\square_1$ | $\square_2$ | $\square_3$ | $\square_4$ 119/ |
| **Your student** failed to show up for a regularly scheduled meeting. | $\square_1$ | $\square_2$ | $\square_3$ | $\square_4$ 120/ |
| **You** failed to show up for a regularly scheduled meeting. | $\square_1$ | $\square_2$ | $\square_3$ | $\square_4$ 121/ |
| It was hard for you to engage your student in conversation. | $\square_1$ | $\square_2$ | $\square_3$ | $\square_4$ 122/ |
| You felt overwhelmed by issues that your student presented you with. | $\square_1$ | $\square_2$ | $\square_3$ | $\square_4$ 123/ |
| Your student seemed bored/disengaged. | $\square_1$ | $\square_2$ | $\square_3$ | $\square_4$ 124/ |
| You enjoyed the time with your student. | $\square_1$ | $\square_2$ | $\square_3$ | $\square_4$ 125/ |
| Your student enjoyed the time with you. | $\square_1$ | $\square_2$ | $\square_3$ | $\square_4$ 126/ |
| Other *(Please specify:)* _____ | $\square_1$ | $\square_2$ | $\square_3$ | $\square_4$ 127/ |

128-129/ 130-131/ 132-133/

**Check one box for each row below …**

17. Did your relationship with your student end over the course of the year?

$\square_1$ Yes
134/
$\square_2$ No → *Go to question 23*

18. If yes, who initiated the end of the match?

$\square_1$ I did.
135/
$\square_2$ The student did. → *Go to question 20*
$\square_3$ The agency or school did. → *Go to question 21*

19. If **you** ended the match, why did you do so? **(Check all that apply.)**

$\square_1$ It was too much of a time commitment
136/
$\square_2$ I didn't get along with my student
137/
$\square_3$ My student had needs that I could not meet; I felt overwhelmed
138/
$\square_4$ The program was disorganized/run poorly
139/
$\square_5$ I did not feel welcome at my student's school
140/

❑₉₅ Other *(Please describe:)*

_____ 141-142/

143-144/ 145-146/ 147-148/

## *If you answered question 19, go to question 21*

20.     If the **student** ended the match, why did he or she do so? **(Check all that apply.)**

❑₁ He or she moved out of the area

149/

❑₂ He or she had a poor attitude towards the match/behaved poorly while we were meeting
and so it                    became impossible for us to work together

150/

❑₃ He or she didn't enjoy the experience

151/

❑₉₅ Other *(Please describe:)*

_____ 152-153/

154-155/ 156-157/ 158-

159/

❑₉₇ I don't know

160-161/

21.     How disappointed were **you** that the match ended?

❑₁ Not disappointed at all

162/

❑₂ A little disappointed
❑₃ Fairly disappointed
❑₄ Extremely disappointed

22.     How disappointed did **your student** appear to be that the match ended?

❑₁ Not disappointed at all

163/

❑₂ A little disappointed
❑₃ Fairly disappointed
❑₄ Extremely disappointed
❑₅ Couldn't tell/not sure

23.     Please record the date you completed this questionnaire.

|___|___|    |___|___|    |___|___|___|___|
  Month        Day            Year
164-165/      166-167/      168-171/

**If you had a mentoring relationship with any other students in this program, whose name appears on one of the attached survey cover pages, please complete the questionnaire for that/those student(s).**

**If you have filled out a questionnaire for each student with whom you have/had a mentoring relationship, the survey is now complete. Please place all completed questionnaires in the pre-paid envelope provided. Thank you for participating!**

*Please return this survey along with your completed Mentor Survey, to:*

**Mentor Survey**
**C/O Abt Associates Inc.**
**55 Wheeler St.**
**Cambridge, MA 02138**

**THANK YOU FOR PARTICIPATING IN THIS SURVEY!**

**Abt**

**Abt Associates Inc.**

## CHILD ASSENT FORM

### Student Mentoring Evaluation

Dear Student:

The U.S. Department of Education gives out money for student mentoring programs. Student mentoring programs are where students meet with older students or adults to help with school, talk about stuff, and hang out. The Department of Education needs a study to find out how helpful these mentoring programs are and how they can be made better. A company called Abt Associates is doing the study. They will be asking students like you about the things they do and how they feel. Some of these students will be part of a mentoring program and some will not. But we need answers to these questions from both kinds of students.

There are no right or wrong answers to these questions. Your answers will be combined with other students' answers. All the answers you give to our questions will be confidential. This means that we will not tell your parents, your teachers, your school, or anyone else who you know about the answers you give us. We will not ask you to put your name on the answer sheet.

You do not have to answer the questions but we hope you will. We will be asking you some questions about things that you do; some of these questions will be about personal things like questions about your family and friends, drug and alcohol use, and how you do in school. If you do not want to answer a question, you may leave it blank. You may stop answering the questions any time you want. You may ask questions to the person giving the survey to you any time you like. Again, all the answers you give to these questions will be private.

We have also obtained a Confidentiality Certificate (CC) from the US Department of Health and Human Services (DHHS) to protect the researchers from being forced, even by court order or subpoena, to identify you. (The Certificate does not imply approval or disapproval of the project by the Secretary of DHHS. It adds special protection for the research information about you.) You should know, however, that researchers may provide information to appropriate individuals or agencies if harm to you, harm to others, or child abuse becomes a concern. In addition, the federal agency funding this research may see your information if it audits us.

Please read the statements below and sign your name, telling us whether or not you will answer the questions. If you do not want to answer the questions nothing bad will happen to you.

PLEASE PUT AN "X" ON ONE OF THE LINES BELOW, AND PRINT AND SIGN YOUR NAME.

_____ YES, I WILL ANSWER THE QUESTIONS IN THIS SURVEY.  My answers will be used for research and will never be given to my parents/guardian, my school or anyone else.

_____ NO, I WILL NOT ANSWER THE QUESTIONS IN THIS SURVEY.


«Name» _____    _____
                                                                                    SIGN YOUR NAME ON THE LINE

# Student Survey
# Fall Version

Abt

**Abt Associates Inc.**

**Student Survey**                                   **Today's Date is:** _____/_____/_____

### DO NOT write your name anywhere on your paper.

Use the pencil provided to complete the survey.
Fill in circles completely.

Like this: ●    Not like this: Ⓙ   Ⓧ   ⊘

---

1.  **You are a . . .**

    ❑  Boy

    ❑  Girl

4.  **Which of these people do you live with most of the time?**
    *(Mark ALL of the people who live with you most of the time, not just now and then.)*

    ❑  Mother                                   ❑  Sister(s)

    ❑  Stepmother                               ❑  Brother(s)

    ❑  Foster mother, female guardian           ❑  Other children

    ❑  Father                                   ❑  Grandparent(s)

    ❑  Stepfather                               ❑  Other adult(s)

    ❑  Foster father, male guardian

---

Please fill in the circle that tells whether these things are Not True at All, Not Very True, Sort of True, or Very True for you.

| 5. About School | | Not True at All | Not Very True | Sort of True | Very True |
|---|---|---|---|---|---|
| 5.a | I'm pretty slow in finishing my schoolwork. | ☐ | ☐ | ☐ | ☐ |
| 5.b | I do well at my classwork. | ☐ | ☐ | ☐ | ☐ |
| 5.c | I have trouble figuring out the answers in school. | ☐ | ☐ | ☐ | ☐ |
| 5.d | I forget what I learn. | ☐ | ☐ | ☐ | ☐ |
| 5.e | I feel that I am just as smart as other kids my age. | ☐ | ☐ | ☐ | ☐ |
| 5.f | I raise my hand in class to answer questions. | ☐ | ☐ | ☐ | ☐ |
| 5.g | I do extra schoolwork on my own. | ☐ | ☐ | ☐ | ☐ |
| 5.h | I feel that I am good at schoolwork. | ☐ | ☐ | ☐ | ☐ |
| 5.i | I like school. | ☐ | ☐ | ☐ | ☐ |
| 5.j | Most mornings I look forward to going to school. | ☐ | ☐ | ☐ | ☐ |
| 5.k | When I have schoolwork to do, I keep working on it until it is finished. | ☐ | ☐ | ☐ | ☐ |

| 6. About Your Friends | | Not True at All | Not Very True | Sort of True | Very True |
|---|---|---|---|---|---|
| 6.a | I find it hard to make friends. | ☐ | ☐ | ☐ | ☐ |
| 6.b | I argue or fight with my friends. | ☐ | ☐ | ☐ | ☐ |
| 6.c | I wish I had more friends. | ☐ | ☐ | ☐ | ☐ |

Please fill in the circle that tells how often you do these things: Never, Not Much, Some, or A Lot.

| 7. About Your Parents/Guardians | | Never | Not Much | Some | A Lot |
|---|---|---|---|---|---|
| 7.a | I talk with my parent(s) about things I do with my friends. | ☐ | ☐ | ☐ | ☐ |
| 7.b | I go to a movie, play, museum, or sports event with my parent(s). | ☐ | ☐ | ☐ | ☐ |
| 7.c | I talk with my parent(s) about a problem I am having. | ☐ | ☐ | ☐ | ☐ |
| 7.d | I talk with my parent(s) about schoolwork or grades. | ☐ | ☐ | ☐ | ☐ |

| 8. About Other Adults | | Never | Not Much | Some | A Lot |
|---|---|---|---|---|---|
| 8.a | I get help with my schoolwork or homework outside of regular school hours from adults other than my parents/guardians. | ☐ | ☐ | ☐ | ☐ |
| 8.b | I get help with problems or stuff that is bothering me from adults other than my parents/guardians. | ☐ | ☐ | ☐ | ☐ |
| 8.c | I hang out or do fun things with adults other than my parents/guardians. | ☐ | ☐ | ☐ | ☐ |

| 9. | About Things You Do | Never | Not Much | Some | A Lot |
|---|---|---|---|---|---|
| 9.a | Finish your homework without being reminded. | ☐ | ☐ | ☐ | ☐ |
| 9.b | Get out of bed for school without your parents or other people having to wake you. | ☐ | ☐ | ☐ | ☐ |
| 9.c | Clean up after yourself without being reminded. | ☐ | ☐ | ☐ | ☐ |
| 9.d | Volunteer to help others through your church, mosque, temple, or synagogue. | ☐ | ☐ | ☐ | ☐ |
| 9.e | Volunteer to help others at your school. | ☐ | ☐ | ☐ | ☐ |
| 9.f | Volunteer to help others in your neighborhood. | ☐ | ☐ | ☐ | ☐ |

| 10. | During the past month, did you . . . | Never | Not Much | Some | A Lot |
|---|---|---|---|---|---|
| 10.a | Break something on purpose? | ☐ | ☐ | ☐ | ☐ |
| 10.b | Punch, kick, or hit someone? | ☐ | ☐ | ☐ | ☐ |
| 10.c | Argue with your parents? | ☐ | ☐ | ☐ | ☐ |
| 10.d | Lie to your parents about something? | ☐ | ☐ | ☐ | ☐ |
| 10.e | Skip school without permission? | ☐ | ☐ | ☐ | ☐ |
| 10.f | Steal something from a store or from another person? | ☐ | ☐ | ☐ | ☐ |
| 10.g | Give a teacher a hard time? | ☐ | ☐ | ☐ | ☐ |
| 10.h | Carry a weapon, such as a club, knife, or gun? | ☐ | ☐ | ☐ | ☐ |

**11.** **Have you ever been a member of a gang?** (A gang is a group that does some illegal things together, and may have a special name or an area it calls its own.)

    ☐   Yes

    ☐   No

**11.a** **Are you now a member of a gang?**

    ☐   Yes

    ☐   No

**12.** **Cigarettes, Alcohol, and Other Drugs**

**12.a** **How many times did you smoke a cigarette or chew tobacco in the past month?**

    ☐   I have never smoked or chewed tobacco

    ☐   None in the last month

    ☐   Once or twice

    ☐   3 to 5 times

    ☐   6 to 9 times

    ☐   10 or more times

**12.b    How many times did you drink a glass of beer, wine, or other alcohol in the past month?**

☐  I have never drunk a glass of          ☐  3 to 5 times
    beer, wine or other alcohol            ☐  6 to 9 times
☐  None in the last month                  ☐  10 or more times
☐  Once or twice

**12.c    How many times in the past month did you use any of the following drugs?**

| | I have never used | None in the past month | Once or twice | 3-5 times | 6-9 times | More than 10 times |
|---|---|---|---|---|---|---|
| Marijuana (pot, weed) | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Glue, gas, aerosol sprays | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Medicine not prescribed for you | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Cocaine, methamphetamine ("speed") | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Some other drug | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |

(Please write its name): _____

**13.    How important is it for you . . .**

| | Not Important at All | Not Very Important | Important | Very Important |
|---|---|---|---|---|
| 13.a    To graduate from high school? | ☐ | ☐ | ☐ | ☐ |
| 13.b    To get an education after finishing high school, such as going to college? | ☐ | ☐ | ☐ | ☐ |
| 13.c    To be successful in a job or career? | ☐ | ☐ | ☐ | ☐ |
| 13.d    To save money for the future? | ☐ | ☐ | ☐ | ☐ |

**14.    Have you ever been in a mentoring program before this school year?**  (By mentoring, we mean a program where you are matched with an adult or older student and you meet to do things, work on schoolwork, or talk about problems.)

☐  Yes
☐  No

**15.    How often did you meet in that program?**

☐  I have never been in a mentoring program before
☐  More than 4 times a month
☐  2 to 4 times a month
☐  Once a month
☐  Less than once a month

**This is the end of our questions. Thank you very much for participating.**

**Abt**

**Abt Associates Inc.**

**CHILD ASSENT FORM**

**Student Mentoring Evaluation**

Dear Student:

The U.S. Department of Education gives out money for student mentoring programs. Student mentoring programs are where students meet with older students or adults to help with school, talk about stuff, and hang out. The Department of Education needs a study to find out how helpful these mentoring programs are and how they can be made better. A company called Abt Associates is doing the study. They will be asking students like you about the things they do and how they feel. Some of these students will be part of a mentoring program and some will not. But we need answers to these questions from both kinds of students.

There are no right or wrong answers to these questions. Your answers will be combined with other students' answers. All the answers you give to our questions will be confidential. This means that we will not tell your parents, your teachers, your school, or anyone else who you know about the answers you give us. We will not ask you to put your name on the answer sheet.

You do not have to answer the questions but we hope you will. We will be asking you some questions about things that you do; some of these questions will be about personal things like questions about your family and friends, drug and alcohol use, and how you do in school. If you do not want to answer a question, you may leave it blank. You may stop answering the questions any time you want. You may ask questions to the person giving the survey to you any time you like. Again, all the answers you give to these questions will be private.

We have also obtained a Confidentiality Certificate (CC) from the US Department of Health and Human Services (DHHS) to protect the researchers from being forced, even by court order or subpoena, to identify you. (The Certificate does not imply approval or disapproval of the project by the Secretary of DHHS. It adds special protection for the research information about you.) You should know, however, that researchers may provide information to appropriate individuals or agencies if harm to you, harm to others, or child abuse becomes a concern. In addition, the federal agency funding this research may see your information if it audits us.

Please read the statements below and sign your name, telling us whether or not you will answer the questions. If you do not want to answer the questions nothing bad will happen to you.

PLEASE PUT AN "X" ON ONE OF THE LINES BELOW, AND PRINT AND SIGN YOUR NAME.

_____ YES, I WILL ANSWER THE QUESTIONS IN THIS SURVEY.  My answers will be used for research and will never be given to my parents/guardian, my school or anyone else.

_____ NO, I WILL NOT ANSWER THE QUESTIONS IN THIS SURVEY.


_____          _____
                                                  SIGN YOUR NAME ON THE LINE

# Student Survey

# Spring Version

The Evaluation of the
Student Mentoring
Program

Abt

Abt Associates Inc.

According to the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless such collection displays a valid OMB control number. The valid OMB control number for this information collection is 1850-0806. The time required to complete this information collection is estimated to average 25 minutes per response, including the time to review instructions, search existing data resources, gather the data needed, and complete and review the information collection. If you have any comments concerning the accuracy of the time estimate(s) or suggestions for improving this form, please write to: U.S. Department of Education, Washington, D.C. 20202-4651. If you have comments or concerns regarding the status of your individual submission of this form, write directly to: Institute of Education Sciences, National Center for Educational Evaluation, U.S. Department of Education, 555 New Jersey Avenue, Room 501, Washington, D.C. 20208.

## DO NOT write your name anywhere on your paper.

Use the pencil provided to complete the survey.
Fill in circles completely.

Like this: ●     Not like this: Ⓥ    Ⓧ    ⊘

1. **You are a . . .**

   ☐  Boy

   ☐  Girl

4. **Which of these people do you live with most of the time?**
   *(Mark ALL of the people who live with you most of the time, not just now and then.)*

   | | |
   |---|---|
   | Mother | Sister(s) |
   | Stepmother | Brother(s) |
   | Foster mother, female guardian | Other children |
   | Father | Grandparent(s) |
   | Stepfather | Other adult(s) |
   | Foster father, male guardian | |

Please fill in the circle that tells whether these things are Not True at All, Not Very True, Sort of True,
or Very True for you.

| 5. | About School | | Not True at All | Not Very True | Sort of True | Very True |
|---|---|---|---|---|---|---|
| 5.a | I'm pretty slow in finishing my schoolwork. | | ☐ | ☐ | ☐ | ☐ |
| 5.b | I do well at my classwork. | | ☐ | ☐ | ☐ | ☐ |
| 5.c | I have trouble figuring out the answers in school. | | ☐ | ☐ | ☐ | ☐ |
| 5.d | I forget what I learn. | | ☐ | ☐ | ☐ | ☐ |
| 5.e | I feel that I am just as smart as other kids my age. | | ☐ | ☐ | ☐ | ☐ |
| 5.f | I raise my hand in class to answer questions. | | ☐ | ☐ | ☐ | ☐ |
| 5.g | I do extra schoolwork on my own. | | ☐ | ☐ | ☐ | ☐ |
| 5.h | I feel that I am good at schoolwork. | | ☐ | ☐ | ☐ | ☐ |
| 5.i | I like school. | | ☐ | ☐ | ☐ | ☐ |
| 5.j | Most mornings I look forward to going to school. | | ☐ | ☐ | ☐ | ☐ |
| 5.k | When I have schoolwork to do, I keep working on it until it is finished. | | ☐ | ☐ | ☐ | ☐ |

| 6. | About Your Friends | | Not True at All | Not Very True | Sort of True | Very True |
|---|---|---|---|---|---|---|
| 6.a | I find it hard to make friends. | | ☐ | ☐ | ☐ | ☐ |
| 6.b | I argue or fight with my friends. | | ☐ | ☐ | ☐ | ☐ |
| 6.c | I wish I had more friends. | | ☐ | ☐ | ☐ | ☐ |

Please fill in the circle that tells how often you do these things: Never, Not Much, Some, or A Lot.

| 7. | About Your Parents/Guardians | | Never | Not Much | Some | A Lot |
|---|---|---|---|---|---|---|
| 7.a | I talk with my parent(s) about things I do with my friends. | | ☐ | ☐ | ☐ | ☐ |
| 7.b | I go to a movie, play, museum, or sports event with my parent(s). | | ☐ | ☐ | ☐ | ☐ |
| 7.c | I talk with my parent(s) about a problem I am having. | | ☐ | ☐ | ☐ | ☐ |
| 7.d | I talk with my parent(s) about schoolwork or grades. | | ☐ | ☐ | ☐ | ☐ |

| 8. | About Other Adults | | Never | Not Much | Some | A Lot |
|---|---|---|---|---|---|---|
| 8.a | I get help with my schoolwork or homework outside of regular school hours from adults other than my parents/guardians. | | ☐ | ☐ | ☐ | ☐ |
| 8.b | I get help with problems or stuff that is bothering me from adults other than my parents/guardians. | | ☐ | ☐ | ☐ | ☐ |
| 8.c | I hang out or do fun things with adults other than my parents/guardians. | | ☐ | ☐ | ☐ | ☐ |

| 9. | About Things You Do | Never | Not Much | Some | A Lot |
|---|---|:---:|:---:|:---:|:---:|
| 9.a | Finish your homework without being reminded. | ☐ | ☐ | ☐ | ☐ |
| 9.b | Get out of bed for school without your parents or other people having to wake you. | ☐ | ☐ | ☐ | ☐ |
| 9.c | Clean up after yourself without being reminded. | ☐ | ☐ | ☐ | ☐ |
| 9.d | Volunteer to help others through your church, mosque, temple, or synagogue. | ☐ | ☐ | ☐ | ☐ |
| 9.e | Volunteer to help others at your school. | ☐ | ☐ | ☐ | ☐ |
| 9.f | Volunteer to help others in your neighborhood. | ☐ | ☐ | ☐ | ☐ |

| 10. | During the past month, did you . . . | Never | Not Much | Some | A Lot |
|---|---|:---:|:---:|:---:|:---:|
| 10.a | Break something on purpose? | ☐ | ☐ | ☐ | ☐ |
| 10.b | Punch, kick, or hit someone? | ☐ | ☐ | ☐ | ☐ |
| 10.c | Argue with your parents? | ☐ | ☐ | ☐ | ☐ |
| 10.d | Lie to your parents about something? | ☐ | ☐ | ☐ | ☐ |
| 10.e | Skip school without permission? | ☐ | ☐ | ☐ | ☐ |
| 10.f | Steal something from a store or from another person? | ☐ | ☐ | ☐ | ☐ |
| 10.g | Give a teacher a hard time? | ☐ | ☐ | ☐ | ☐ |
| 10.h | Carry a weapon, such as a club, knife, or gun? | ☐ | ☐ | ☐ | ☐ |

**11. Have you ever been a member of a gang?** (A gang is a group that does some illegal things together, and may have a special name or an area it calls its own.)

☐  Yes

☐  No

**11.a  Are you now a member of a gang?**

☐  Yes

☐  No

**12.  Cigarettes, Alcohol, and Other Drugs**

**12.a  How many times did you smoke a cigarette or chew tobacco in the past month?**

☐  I have never smoked or chewed tobacco

☐  None in the last month

☐  Once or twice

☐  3 to 5 times

☐  6 to 9 times

☐  10 or more times

**12.b** **How many times did you drink a glass of beer, wine, or other alcohol in the past month?**

☐ I have never drunk a glass of beer, wine or other alcohol

☐ None in the last month

☐ Once or twice

☐ 3 to 5 times

☐ 6 to 9 times

☐ 10 or more times

**12.c** **How many times in the past month did you use any of the following drugs?**

| | I have never used | None in the past month | Once or twice | 3-5 times | 6-9 times | More than 10 times |
|---|---|---|---|---|---|---|
| Marijuana (pot, weed) | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Glue, gas, aerosol sprays | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Medicine not prescribed for you | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Cocaine, methamphetamine ("speed") | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| Some other drug (Please write its name): _____ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |

**13.** **How important is it for you . . .**

| | Not Important at All | Not Very Important | Important | Very Important |
|---|---|---|---|---|
| 13.a To graduate from high school? | ☐ | ☐ | ☐ | ☐ |
| 13.b To get an education after finishing high school, such as going to college? | ☐ | ☐ | ☐ | ☐ |
| 13.c To be successful in a job or career? | ☐ | ☐ | ☐ | ☐ |
| 13.d To save money for the future? | ☐ | ☐ | ☐ | ☐ |

**14.** **Have you been involved in a mentoring program this past school year?** (By mentoring we mean a program where you are matched with an adult or older student and you meet to do things, work on schoolwork, or talk about problems.)

☐ Yes

☐ No

*If you answered No to this question, you are finished with the survey. Please put down your pencil and wait while we finish with others.*

**15.** **How often do you meet with your mentor in that program?**

☐ I have never been in a mentoring program before

☐ More than 4 times a month

☐ 2 to 4 times a month

☐ Once a month

☐ Less than once a month

**16.** **Were you mentored as part of the _____ program this past school year?**

☐ Yes

☐ No

*If you answered No to this question, you are finished with the survey. Please put down your pencil and wait while we finish with others.*

These last few questions are for only those of you were mentored as part of the _____ program. We will be asking you about the person who has been assigned to you by that mentoring program, your mentor. Please answer these questions about you and your mentor.

**17.** **Since school started in fall, how many different mentors from this program have you met with?**

☐ None

☐ One

☐ Two

☐ More than two

*If you answered **None** to this question, you are finished with the survey. Please put down your pencil and wait while we finish with others.*

**18.** **Are you still meeting with a mentor?**

☐ Yes *(Go to question 19)*

☐ No

**18.a** *(If No:)* **How long ago did you stop meeting with your mentor?**

☐ This month

☐ Last month

☐ Between last month and six months ago

☐ More than six months ago

Please fill in the circle that tells whether these things are Not True at All, Not Very True, Sort of True,
or Very True for you.

**19.**

| Please answer these questions about you and your mentor. | Not True at All | Not Very True | Sort of True | Very True |
|---|---|---|---|---|
| 19.a When I am with my mentor, I feel bored. | ☐ | ☐ | ☐ | ☐ |
| 19.b Sometimes my mentor promises that we will do something and then we don't do it. | ☐ | ☐ | ☐ | ☐ |
| 19.c I feel that I can trust my mentor. | ☐ | ☐ | ☐ | ☐ |
| 19.d When something is bugging me, my mentor listens to me. | ☐ | ☐ | ☐ | ☐ |
| 19.e My mentor has good ideas about how to solve problems. | ☐ | ☐ | ☐ | ☐ |
| 19.f My mentor talks to me about my future. | ☐ | ☐ | ☐ | ☐ |
| 19.g My mentor helps me with my schoolwork. | ☐ | ☐ | ☐ | ☐ |

**This is the end of our questions. Thank you very much for participating.**

# Appendix C: Construction of Student Outcome Measures

As explained in Chapter 2, outcome measures for the Impact Study were derived from two sources: the Student Survey and school records. The purpose of this appendix is to explain in detail the creation of the outcome measures based on items in the Student Survey and from abstraction of school records. Readers may refer to Exhibit 2.4, which summarizes the sources and measures used in developing student outcome measures, as well as Appendix B for copies of all survey instruments.

## Construction of Measures Based on Student Survey Data

The scales included in the Student Survey were selected with the program logic model in mind; that is, specific measures of impact were matched to the specific goals and activities of the mentoring program. The scales selected met requirements of adequate reliability and validity and had been used in similar studies. Exhibit C.1 identifies the original scales and items in the Student Survey representing the impact domains of interest.

**Exhibit C.1**

**Student Survey—Included Scales and Measures**

| Scale/Measure | Impact Domain | Question Number | Source |
|---|---|---|---|
| Gender | N/A | 1 | US Census |
| Household Composition | N/A | 4 | Original Measure |
| Scholastic Efficacy | Academics | 5a-f* | Subscale of Harter (1988) *Self Perception Profile for Adolescents* <br> *5f is original item |
| School Bonding | Academics | 5g-k | Hawkins, Guo, Hill, Battin-Pearson, and Abbot (2001) |
| Peer Relationships | Pro-Social Behaviors | 6a-c | Original Scale |
| Parental Relationships | Pro-Social Behaviors | 7a-d | Adapted from *National Longitudinal Study of Adolescent Health (AddHealth)* – National Institute of Child Health and Human Development |
| Relationships with Other Adults | Pro-Social Behaviors | 8a-c | Original Scale |
| Personal Initiative | Pro-Social Behaviors | 9a-f | Modified from *Michigan State University Early Adolescent Survey II*, Michigan State University Cooperative Extension Service |
| Delinquency/Misconduct | Delinquency/ Misconduct | 10a-h | Adapted from *21st Century Community Learning Centers Program* survey, Mathematica Policy Research |
| Gang Involvement | Delinquency/ Misconduct | 11 | Original Measure |
| Tobacco Use | Delinquency/ Misconduct | 12a | Adapted from *Monitoring the Future*, SAMHSA |
| Alcohol Use | Delinquency/ Misconduct | 12b | Adapted from *Monitoring the Future*, SAMHSA |
| Drug Use | Delinquency/ Misconduct | 12c | Adapted from *Monitoring the Future*, SAMHSA |
| Future Orientation | Academics | 13a-d | Original Scale |

The Student Survey contains a number of scales representing each domain. In developing final outcome measures from the survey items we undertook a series of steps to refine and confirm the best scale construction for our sample. We took these steps for several reasons. First, most of the scales originally identified for inclusion in the Student Survey were subsequently altered in some way, e.g. items were added or subtracted, or subscales of larger scales were used. In addition, the ultimate study sample spanned a large age and comprehension range, from 4th to 8th grade, and some of the scales selected were originally developed on somewhat older populations of children. Given these concerns, we had two analytic goals in constructing final impact measures from the Student Survey: 1) to confirm the reasonableness of the items that constitute the scales originally selected for the instrument, and 2) to determine the reliability of those scales in our sample, eliminating those with low reliability coefficients.

### Initial Factor Analyses

We performed a series of factor analyses to determine the utility of our outcome measures for the populations we surveyed. We originally planned a traditional "confirmation" of the underlying constructs represented in the subscales that had been taken from other instruments for use in this context. We were alert to the fact that the subscales had been extracted from longer instruments and in some cases wording had been changed and single items added. In addition, we were aware that while some of the scales have sound psychometric properties they were not tested on populations as diverse in age as the children in our study. In the best of all cases, the factor analysis would have simply confirmed the 8 dimensions they were thought to measure. This was not the case. We describe the steps taken in the development of final outcome measures below.

Assuming that we may be confirming existing scale constructions, we first analyzed sets of scales or items that appeared together visually under the same topic heading on the survey (e.g., *About School*) and/or constituted existing scales.[1] For example, the original 10 items under *About School* constituted subsets of items from two previously developed scales: Scholastic Efficacy (Harter, 1988) and School Engagement/Bonding (Hawkins et al., 2001), with one original item added. We examined each of the similar blended areas of the survey without specifying the number of factors to be extracted within each, assuming the original scales would emerge intact. Scales were analyzed using the fall data. Again, because our survey questions were developed based on existing instruments that have been previously tested and validated by other researchers, our working hypothesis was that the factors present would correspond to these instruments, and that individual items would load in groupings consistent with the original survey question structure, for a total of eight factors in all.

However, at the end of the initial factor analyses, there were 11 preliminary groupings:

- School Bonding
- Scholastic Efficacy
- Perceived Learning Difficulty
- Volunteerism
- Personal Responsibility
- Misconduct
- Delinquency

---

[1]  These preliminary analyses did not include question 11 (Gang Involvement) or question 12 (tobacco, alcohol, and drug abuse).

- Peer Relationships
- Relationships with Parents
- Relationship with Other Adults
- Future Orientation

We evaluated the reliability of each of these outcome measures by calculating Cronbach's $\alpha$, a measure of internal consistency, using a reliability cutoff of $\alpha = .70$ as a rule of thumb.[2] Only three groupings, school bonding, future orientation, and mentoring relationships, had Cronbach's $\alpha$ above .70, and four groupings (school efficacy, relationship with parents, relationship with other adults, and misconduct) had Cronbach's $\alpha$ above .60 but below .70.

At this point there were several possible options available to us if we wished to retain these natural groupings: 1) dropping groupings with low reliability (relative to either the .60 or .70 cutoff), 2) retaining groupings with Cronbach's $\alpha$ above 0.60 as well as groupings with Cronbach's $\alpha$ relatively close to the .60 cutoff on the basis of their strong theoretical validity, and/or 3) choosing individual items of interest as single-item outcome measures. We judged the first option to be too restrictive, because it would have eliminated too many groupings, sacrificing important information from each impact domain. The second option was deemed to unacceptably compromise reliability standards. The third and final option threatened analytic parsimony, and additionally could have lent the appearance of "cherry-picking" individual items to achieve desired results.

**Exploratory Factor Analysis**

We therefore undertook an exploratory factor analysis using all of the survey items to determine the most efficient and reliable set of constructs for this sample and to ascertain whether any other natural groupings of survey items across questions might be present, in hopes of improving both the range of impact measures and their reliability. The new analysis also included imputed data for scales with missing items.[3]

In general, the results were quite similar to the results from the first analysis, with items falling into the same groupings with only a few minor adjustments involving the retention or exclusion of individual items from each scale. In an effort to improve the overall reliability of other factor groupings to allow the inclusion of these impact measures, we continued the analysis including items from questions 11 and 12, which asked about gang membership and the use of illegal substances (alcohol, tobacco, and drugs), respectively. The result was essentially identical to the prior results, with gang membership and illegal substance abuse items grouping to form two additional outcome measures, for a total of 13 individual factors, only four of which had Cronbach's $\alpha$ above .70: (School Bonding, Future Orientation, Gang Membership, and Misconduct).

---

[2] However, particularly for scales with a small number of items, a Cronbach's $\alpha$ of .60 may in some instances be considered acceptable (DeVellis, 2003).

[3] These values were imputed according to the following rules:
- On scales (factors) that are made up of five items or more, allow for one missing item and impute it from the mean. Do not include observations with more than one missing item.
- On scales with fewer than five items, only include observations with all values present.
98% of observations were included based on these criteria.

**Restricted Analyses**

Finally, we conducted a series of *restricted analyses*, varying the power of the analysis and specification of the number of factors allowed. A seven-factor solution resulted in constructs with acceptable theoretical validity for all seven factors. All but two factors (Self-Perception (.61) and Relationship with Other Adults (.61)) were also viewed as having acceptable reliability. Four of these had Cronbach's α at or above the .70 cutoff, while one (Pro-social Behaviors) had an α of .69.[4] Consequently, we eliminated the two factors with unacceptable reliability, resulting in a total of five outcome measures from the Student Survey scales as follows:

- Pro-social Behaviors (combines items on parental relationships, volunteerism, and personal responsibility)
- Delinquency (combines general items on delinquent behaviors with specific items on gang membership and alcohol, drug, and tobacco use)
- Misconduct
- Scholastic Efficacy and School Bonding (combines items from the Scholastic Efficacy scale and the School Bonding scale)
- Future Orientation

Exhibit C.2 reports included survey items and reliability coefficients for each scale.

---

**Exhibit C.2**

**Outcome Measures and Reliability Coefficients: Student Survey**

| Measure | Cronbach's Alpha |
|---|---|
| *Pro-social Behaviors* | .69 |
| I talk with my parent(s) about things I do with my friends | |
| I go to a movie, play, museum or sports event with my parent(s). | |
| I talk to my parents about a problem | |
| I talk to my parents about schoolwork | |
| Finish your homework without being reminded | |
| Get out of bed for school without being reminded | |
| Clean up after yourself without being reminded | |
| Volunteer to help others through church, mosque, temple or synagogue | |
| Volunteer to help others at school | |
| Volunteer to help others in your neighborhood | |
| *Future Orientation* | .76 |
| How important is it for you: | |
| To graduate from high school | |
| To get an education after finishing high school, such as going to college? | |
| To be successful in a job or career? | |

---

[4]   Given the direct relevance of this domain to the mentoring intervention and its close proximity to the cutoff, we felt it was acceptable to include it in our measures.

---

**Exhibit C.2**

**Outcome Measures and Reliability Coefficients: Student Survey**

| Measure | Cronbach's Alpha |
|---|---|
| ***Misconduct*** | .72 |
| *(Reverse coded)* During the past month did you: | |
| Break something on purpose | |
| Punch, kick or hit someone | |
| Argue with your parents | |
| Lie to your parents about something | |
| Steal something from a store or another person | |
| Give a teacher a hard time | |
| Argue or fight with my friends | |
| ***Delinquency*** | .74 |
| *(Reverse coded)* During the past month did you: | |
| Skip school without permission | |
| Carry a weapon, such as a club, knife or gun | |
| Have you ever been a member of a gang? | |
| Are you now a member of a gang? | |
| How many times did you smoke a cigarette or chew tobacco in the past month? | |
| How many times did you drink a glass of beer, wine or other alcohol in the past month? | |
| How many times in the past month did you use any of the following drugs? *(Followed by list grid of frequency of marijuana; glue, gas aerosols; cocaine, methamphetamine; some other drug (specified))* | |
| ***Scholastic Efficacy & School Bonding*** | .72 |
| I do well at my classwork | |
| I feel that I am just as smart as other kids my age | |
| I raise my hand in class to answer questions | |
| I do extra schoolwork on my own | |
| I feel that I am good at schoolwork | |
| I like school | |
| Most mornings I look forward to going to school | |
| When I have schoolwork to do, I keep working on it until it is finished | |

All items except for several of the Delinquency measures were answered using the same 1 – 4 Likert scale. Four items in the Delinquency measure were scored using different metrics .The drug, alcohol and tobacco items were answered in terms of the frequency of use in 6 ordinal categories from "never used" and "none" to "10 or more times" covering the prior 30 days period. The gang involvement items were dichotomous (i.e., currently in a gang or not). All items were standardized into a common scale for analysis. All of these items were standardized to the mean and standard deviation of the two Likert-scale Delinquency measures.

For example, for a given item $Y_i$, the standardized value $Y_i'$ is given by:

$$[1] \quad Y_i' = \frac{S*(Y_i - \overline{Y})}{s_Y} + M \, ,$$

where,

> $S$ = standard deviation of Likert-scale items,
>
> $M$ = mean of Likert-scale items,
>
> $\overline{Y}$ = mean of $Y_i$, and
>
> $s_Y$ = standard deviation of $Y_i$.

## Construction of Measures Based on School Records

This section explains how final outcome measures were constructed from school records. School records were abstracted for statewide proficiency test scores, grades, attendance, truancy, disciplinary actions, and student demographic factors such as receipt of free or reduced-price school lunch, for the year prior to the study and at the end of the study year.

### Statewide Assessments:  Proficiency Test Scores

The test scores for each site varied across all sites that provided test data (see Exhibit C.3); scores were converted into a dichotomous variable representing the threshold level of what was deemed by that state's standard (as published on state education websites and/or determined via telephone follow up with state education representatives) as "proficient."

**Exhibit C.3: Coding Rules for Converting Statewide Assessment Test Scores to Proficiency Levels by Site**

| Site # | Coded as Proficient | Coded as Not Proficient |
|---|---|---|
| 1 | Score of 300 and higher | Scores below 300 |
| 2 | Score of 300+ or 800+ depending on grade and subject | Scores below 300 or below 800 depending on grade and subject |
| 3 | Proficient (P) & Advanced (A) | Needs Improvement (NI) & Warning/Failing (W) |
| 4 | Achievement levels 3 & 4 | Achievement levels 1 & 2 |
| 5 | "Pass" or score of 2100+ | "Fail" or score below 2100 |
| 6 | A "grade equivalent" score higher than a student's "grade placement" | A "grade equivalent" score lower than a student's "grade placement" |
| 7 | Score of 2100 and higher | Score below 2100 |
| 8 | Meets Standard (M) & Exceeds (E) | Falls Far Below Standard (FFB) & Approaches Standard (A) |
| 9 | Proficient (P) & Distinguished (D) | Novice (N) & Apprentice (A) |
| 10 | Proficient (P), Accelerated (AC) & Advanced (AD) | Limited (L) & Basic (B) |
| 11 | *No data available* | *No data available* |
| 12 | Meets Standard (M) & Exceeds (E) | Academic Warning (W) & Below Standards (B) |
| 13 | Mastery (M) & Advanced (A) or Foundational (F) | Unsatisfactory (U), Approaching Basic (AB) & Basic (B) or Pre-Foundational (PF) |
| 14 | Proficient (3) & Advanced (4) | Minimal (1) & Basic (2) |
| 15 | A "grade equivalent" score higher than a student's "grade placement" | A "grade equivalent" score lower than a student's "grade placement" |
| 16 | Levels 3, 4, & 5 or Meets Standard (M) & Exceeds Standard (E) | Levels 1 & 2 or Partially Meets Standard (P) & Does Not Meet Standard (D) |
| 17 | Meets Standard or scores ending in 50 and higher | Partially Meets Standard or scores ending in less than 50 |
| 18 | Proficient (P) & Advanced (A) | Novice (N) & Partially Proficient (PP) |
| 19 | Achievement levels 3 & 4 | Achievement levels 1 & 2 |
| 20 | Achievement levels 3 & 4 | Achievement levels 1 & 2 |
| 21 | Score of 2100 and higher or Pass | Scores below 2100 or Did Not Pass |
| 22 | Score of 300+ or 800+ depending on grade and subject | Scores below 300 or below 800 depending on grade and subject |
| 23 | Proficiency levels 3 & 4 | Proficiency levels 1 & 2 |
| 24 | Score of 2100 and higher or Pass | Scores below 2100 or Fail |
| 25 | Pass | Fail |
| 26 | Proficient (4) & Advanced (5) | Far Below Basic (1), Below Basic (2), & Basic (3) |
| 27 | Cut off score varies based on grade & subject (ranges from 204-231) | Cut off score varies based on grade & subject     (ranges from 204-231) |
| 28 | Score of 2100 and higher | Scores below 2100 |
| 29 | Proficiency levels 3 & 4 | Proficiency levels 1 & 2 |
| 30 | Proficiency levels 3 & 4 | Proficiency levels 1 & 2 |
| 31 | Meets Standard (M) & Exceeds (E) | Conditional (C) & Do Not Meet (D) |
| 32 | Meets Standard (M) & Exceeds (E) | Academic Warning (W or A) & Below average (B) |

Exhibit C

C-7

## Student Grades

Grades were abstracted using the grading system employed in each school/district and then transformed into a 5-point scale. Grading systems were remarkably similar. Sixty-nine percent of all sites used a single letter in a series (i.e., A-F) or single number in an ordinal series (1-5) system; 31 percent used an interval numeric system (0-100) and less than one percent used some other system (checks, +/-, written text). The three most common grading systems were transformed into grade equivalencies as outlined in Exhibit C.4. In the case of grading systems not listed in the exhibit, we contacted the local school district to determine appropriate conversion rules.

**Exhibit C.4**

**Grade Equivalencies Across Sites**

| Performance Level | Grade Equivalencies Included in Score | |
| --- | --- | --- |
| | 0-100 Scale | Letter Value |
| 1 | 1-59 | F |
| 2 | 60-69 | D |
| 3 | 70-79 | C |
| 4 | 80-89 | B |
| 5 | 90-100 | A |

## Disciplinary Infractions

A wide range of behaviors was reported in school records as reportable infractions. As Exhibit C.5 indicates, these infractions were first categorized into seven broad categories, and counts were developed for the number of each type of infraction reported for each child. These seven categories were subsequently combined to represent less and more serious behaviors. The categories designated as Harassment, Non-Compliance, Truancy, and "other" were combined to create a single "Misconduct" measure. The categories designated as Property, Drug-Related, and Violence were combined to create a single "Delinquency" measure.

**Exhibit C.5**

**Disciplinary Infractions, by Category**

*Harassment:*
    Harassment/Bullying (unspecified)
    Verbal Abuse/Name Calling
    Discrimination
    Indecent Exposure
    Makes an Unfounded Charge against Authority

*Non-Compliance:*
    Inappropriate Behavior/Language, Disruptive
    Failure to Comply with Rules/Disrespectful
    Unauthorized Use of Technology
    Lying/False Information
    Lacks Assignment or Materials
    Bus Violation
    Cheating
    Dress Code Violation
    Not Paying Attention
    Safety Violation
    Student Attire
    Too Many Referrals

*Property:*
    Damage/Deface Property
    Theft/Extortion
    Record Falsification/Tampering
    Fires/False Alarms
    Unauthorized Use of School Property

*Drug-Related:*
    Alcohol
    Drugs
    Prohibited Substance
    Tobacco Use

*Truancy:*
    Cutting Class/Assembly
    Excessive Tardiness
    Leaving School or Class without Permission
    Failure to Attend/Serve Detention
    Chronic Truancy

*Violence:*
    Physical Aggression – Student
    Physical Aggression – Teacher/Staff
    Possession of Dangerous Weapon
    Possession of Explosives/Pyrotechnic Device
    Sexual Misconduct/Assault
    Fighting/Assault (unspecified)
    Endangered Self or Others
    Gang/Cult Activity
    Hazing
    Throwing Objects
    Violent Pictures/Usage of Weapons
    Possession of Other Weapon, No Intent

*Other infractions* (not fitting into other categories)

# Appendix D: Impact Analysis Results on Original Student Survey Scales and Measures

As first discussed in Chapter 2 and elaborated upon in Appendix C, most measures included in the Student Survey were derived from existing scales with adequate levels of reliability and validity previously established in prior research. However, because some of these scales were developed on older populations of students than those represented by our study sample, most were altered in some way prior to their inclusion in the survey instrument. We ultimately found that for our study sample, many of these Student Survey scales did not meet standard minimal criteria for internal reliability.

To correct for potential threats to internal reliability, in developing final outcome measures from Student Survey data we performed principal components factor analysis with Promax rotation to refine and confirm scale construction for our sample. Appendix C describes in detail the factor analytic steps taken in developing the final Student Survey outcome measures, which maximize internal reliability while preserving a logical mapping to the impact domains of interest. However, analysis of these composite scale outcomes could potentially mask meaningful variation in the individual measures making up each scale, and readers may be interested in comparing the results of this study to impacts on scales corresponding to those used in prior research. For these reasons, in this appendix we present results of a supplementary impact analysis conducted using scales as they originally appeared in the Student Survey.[1] We dropped individual items from scales as necessary to improve internal reliability, but otherwise present results based on intact measures representing the original intent of the Student Survey instrument. Exhibit D.1 displays the individual items comprising each of these measures, along with estimated scale reliability and sources for each set of question items.

---

[1] See Appendix B for a copy of the Student Survey instrument.

**Exhibit D.1: Original Student Survey Scales and Measures**

| Scale/ Measure | Impact Domain | Cronbach's Alpha (Fall) | Question Number | Item | Source |
|---|---|---|---|---|---|
| Scholastic Efficacy | Academics | 0.55 | 5.a | I am slow in finishing school work | Subscale of Harter (1988) *Self Perception Profile for Adolescents* |
| | | | 5.b | I do well at class work | |
| | | | 5.c | I have trouble figuring out the answers | |
| | | | 5.d | I forget what I learn | |
| | | | 5.e | I am just as smart as other kids my age | |
| | | | 5.f | *I raise my hand in class | *Original Item |
| School Bonding | Academics | 0.66 | 5.g | I do extra schoolwork on my own | Hawkins, Guo, Hill, Battin-Pearson, and Abbot (2001) |
| | | | 5.h | I am good at school work | |
| | | | 5.i | I like school | |
| | | | 5.j | I look forward to going to school | |
| | | | 5.k | I keep working on schoolwork until it is finished | |
| Peer Relationships | Interpersonal Relationships | 0.56 | 6.a | I find it hard to make friends | Original Scale |
| | | | 6.b | ~~I argue or fight with my friends~~ | |
| | | | 6.c | I wish I had more friends | |
| Parental Relationships | Interpersonal Relationships | 0.63 | 7.a | I talk with my parents about things | Adapted from *National Longitudinal Study of Adolescent Health (AddHealth)* - National Institute of Child Health and Human Development |
| | | | 7.b | ~~I go to a movie with parents~~ | |
| | | | 7.c | I talk with my parents about a problem | |
| | | | 7.d | I talk to my parents about schoolwork | |
| Relationships with Other Adults | Interpersonal Relationships | 0.61 | 8.a | I get help with my schoolwork (other than parents) | Original Scale |
| | | | 8.b | I get help with problems (other than parents) | |
| | | | 8.c | I hang out (other than parents) | |

**Exhibit D.1: Original Student Survey Scales and Measures**

| Scale/ Measure | Impact Domain | Cronbach's Alpha (Fall) | Question Number | Item | Source |
|---|---|---|---|---|---|
| Personal Initiative | Delinquency/ Misconduct | 0.55 | 9.a | Finish your homework w/o being reminded | Modified from *Michigan State University Early Adolescent Survey II*, Michigan State University Cooperative Extension Service |
| | | | 9.b | Get out of bed w/o others waking you | |
| | | | 9.c | Clean up after yourself w/o being reminded | |
| | | | 9.d | Volunteer to help others through church | |
| | | | 9.e | Volunteer to help others at school | |
| | | | 9.f | Volunteer to help others in neighborhood | |
| Delinquency/ Misconduct | Delinquency/Misconduct | 0.72 | 10.a | Break something on purpose | Adapted from *21st Century Community Learning Centers Program* survey, Mathematica Policy Research |
| | | | 10.b | Punch, kick or hit someone | |
| | | | 10.c | Argue with your parents | |
| | | | 10.d | Lie to your parents about something | |
| | | | 10.e | Skip school without permission | |
| | | | 10.f | Steal something from a store or another person | |
| | | | 10.g | Give teacher a hard time | |
| | | | 10.h | Carry a weapon | |
| Gang Involvement | Delinquency/Misconduct | N/A | 11 | Ever/now a member of a gang | Original Measure |
| Tobacco Use | Delinquency/Misconduct | N/A | 12.a | Cigarette consumption | Adapted from *Monitoring the Future*, SAMHSA |
| Alcohol Use | Delinquency/Misconduct | N/A | 12.b | Alcohol consumption | Adapted from *Monitoring the Future*, SAMHSA |
| Drug Use | Delinquency/Misconduct | N/A | 12.c | Drug consumption | Adapted from *Monitoring the Future*, SAMHSA |
| Future Orientation | Academics | 0.76 | 13.a | How important is it: graduate HS | Original Scale |
| | | | 13.b | How important is it: education after HS | |
| | | | 13.c | How important is it: success in career | |
| | | | ~~13.d~~ | ~~How important is it: save money for future~~ | |

Survey items that have been ~~crossed out~~ were dropped from composite scales to maximize internal reliability.

## Scale Reliability

Recall that Cronbach's $\alpha$ is a measure of internal reliability. The reader is cautioned to note that, as is evident from Exhibit D.1, many of the original scales did not meet standard minimal criteria for internal reliability. Only two of eight scales had Cronbach's $\alpha$ of above 0.70 based on Student Survey data collected at baseline; and three of eight had Cronbach's $\alpha$ less than 0.60. Analysis of unreliable outcome data increases the likelihood that random noise in the data will bias impact estimates. **Therefore, results on scales with low reliability reported in this section must therefore be considered merely illustrative, supplemental to the impacts presented in the main text.**

Of additional note is the fact that, for seven of eight Student Survey scales, internal reliability increased significantly between baseline data collected in the fall survey, and post-treatment data collected in the spring. In the spring, four of the eight Student Survey scales had Cronbach's $\alpha$ above 0.70, and three more had Cronbach's $\alpha$ between 0.60 and 0.70.

## Results

Exhibits D.2-D.4 present the results of our impact analyses for the full student sample based on the original Student Survey outcomes. Exhibits D.5-D.19 present results for student subgroups (boys versus girls, students below age 12 versus students aged 12 and up, students from two-parent families versus students from other family structures, students with self-reported delinquent behaviors at baseline versus students with no self-reported delinquent behaviors at baseline, and students who were proficient in both math and reading/English language arts at baseline versus students who were not).

Results are presented by outcome domain in a format parallel to that for our main findings in Chapter 4, so that the reader may easily compare the two sets of findings. Also as in our main impact findings, we performed the Benjamini-Hochberg correction to control for multiple comparisons, as described in Section 2.9. Results for data abstracted from student records are therefore presented along with the Student Survey outcomes in order to define the appropriate "families" of comparisons across which we wished to adjust.

In general, this alternative approach to defining Student Survey outcome measures did not substantively alter the findings of this study. For the full sample, there were no statistically significant impacts of the Student Mentoring Program on pro-social behaviors, academic engagement and achievement, or participation in high-risk behavior or delinquency.

For our subgroup findings, impacts of the Student Mentoring Program on two measures, Scholastic Efficacy and School Bonding, were positive and statistically significant for girls (but not for boys); differences between girls and boys on these measures were also statistically significant. Additionally, there was a statistically significant impact on Future Orientation for boys, though the differences in impacts between girls and boys on this measure was not statistically significant.

There were a few scattered statistically significant findings in the remainder of our subgroup analyses, though there were no statistically significant differences between subgroups other than the gender differences described above. In particular, the truancy rate was statistically significantly lower in treatment group students below the age of 12 relative to their control group counterparts, but not for

the older student group. There was a statistically significant impact on the Relationships with Other Adults scale for students who were not academically proficient in both math and reading/ELA at baseline, but not for proficient students. Finally, there was a statistically significant impact on Scholastic Efficacy for students who reported any delinquent behaviors at baseline, but not for students without any delinquent behaviors.

**Exhibit D.2: Estimated Impact on Interpersonal Relationships, Personal Responsibility, and Community Involvement**

| Self-Reported Outcomes | Unadjusted Mean Outcome | | | | Estimated Impact | | | |
|---|---|---|---|---|---|---|---|---|
| | Treatment Group | | Control Group | | Regression Adjusted T-C Group Difference[b] | p-value to Test Difference | BH-Corrected Critical Value[c] | Estimated Effect Size |
| | Mean | Standard Deviation[a] | Mean | Standard Deviation | | | | |
| Peer Relationships | 2.98 | 1.31 | 2.97 | 1.35 | 0.03 | 0.39 | 0.03 | 0.03 |
| Parental Relationships | 2.95 | 1.13 | 2.94 | 1.12 | 0.01 | 0.79 | 0.05 | 0.01 |
| Relationships with Other Adults | 2.28 | 1.18 | 2.18 | 1.15 | 0.08* | 0.02 | 0.01 | 0.09 |
| Personal Initiative | 2.74 | 0.96 | 2.75 | 0.91 | -0.01 | 0.66 | 0.04 | -0.02 |
| Number of students | 1163 | 1197 | | | | | | |
| Percent missing data | ≤2% | ≤2% | | | | | | |

[a]   Standard Deviations are only reported for Means or Mean Percents.

[b]   Regression Adjusted T-C Difference will not necessarily be equal to the difference between the Unadjusted Mean Outcomes.

[c]   Based on Benjamini-Hochberg test.

*   p-value (of adjusted difference in means) < 0.05, two-tailed test.

+   p-value (of adjusted difference in means) < BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005–Spring 2006; Fall 2006–Spring 2007.

**Exhibit D.3: Estimated Impact on Attitudinal and Academic Outcomes**

| Self-Reported Attitudinal Outcomes (Range 1 – 4) | Unadjusted Mean Outcome | | | | Estimated Impact | | | |
|---|---|---|---|---|---|---|---|---|
| | Treatment Group | | Control Group | | Regression Adjusted T-C Group Difference[b] | p-value to Test Difference | BH-Corrected Critical Value[c] | Estimated Effect Size |
| | Mean | Standard Deviation[a] | Mean | Standard Deviation | | | | |
| Scholastic Efficacy | 3.02 | 0.76 | 2.97 | 0.80 | 0.05* | 0.02 | 0.01 | 0.09 |
| School Bonding | 2.96 | 0.93 | 2.94 | 1.00 | 0.02 | 0.37 | 0.03 | 0.03 |
| Future Orientation | 3.85 | 0.54 | 3.80 | 0.63 | 0.03* | 0.04 | 0.02 | 0.08 |
| Number of students | 1163 | | 1197 | | | | | |
| Percent missing data | ≤2% | | ≤3% | | | | | |
| **School-Reported Academic Outcomes** | | | | | | | | |
| *Overall Absenteeism Rate (Percent)* [d] | 5.03 | 7.71 | 5.49 | 9.63 | -0.46* | 0.04 | 0.01 | -0.09 |
| Number of students | 1163 | | 1197 | | | | | |
| Percent missing data | 15% | | 18% | | | | | |
| *Grades (Range 1–5)* [e] | | | | | | | | |
| Math | 3.19 | 1.70 | 3.23 | 1.67 | -0.05 | 0.23 | 0.02 | -0.05 |
| English Language Arts | 3.57 | 1.78 | 3.61 | 1.69 | -0.04 | 0.40 | 0.04 | -0.04 |
| Science | 3.52 | 1.87 | 3.55 | 1.86 | -0.03 | 0.48 | 0.05 | -0.03 |
| Social Studies | 3.53 | 1.92 | 3.56 | 1.83 | -0.01 | 0.78 | 0.05 | -0.01 |
| Number of students | 1163 | | 1197 | | | | | |
| Percent missing data | ≤35% | | ≤33% | | | | | |
| *State Assessment Tests* | | | | | | | | |
| Math—Percent Proficient | 45.69 | | 47.10 | | -1.53 | 0.41 | 0.04 | 0.94 [f] |
| Reading/ELA—Percent Proficient | 49.40 | | 50.76 | | -1.66 | 0.37 | 0.03 | 0.94 [f] |
| Number of students | 1163 | | 1197 | | | | | |
| Percent missing data | ≤25% | | ≤21% | | | | | |

[a] Standard Deviations are only reported for Means or Mean Percents.

[b] Regression Adjusted T-C Difference will not necessarily be equal to the difference between the Unadjusted Mean Outcomes.

[c] Based on Benjamini-Hochberg test.; figure shown provides the critical value that the "p-value to Test Difference" in the preceding column must be less than in order for the "Regression Adjusted T-C Group Difference" to be statistically significant after controlling for multiple tests.

[d] Lower Overall Absenteeism Rates indicate more positive outcomes.

[e] Higher scores indicate higher grades; see Appendix F for further explanation of how these scores were derived.

[f] Odds-ratio.

\* p-value (of adjusted difference in means) < 0.05, two-tailed test.

\+ p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit D.4: Estimated Impact on Delinquent Behaviors and Participation in Harmful Activities**

| Self-Reported Behavioral Outcomes | Unadjusted Mean Outcome | | | | Estimated Impact | | | |
|---|---|---|---|---|---|---|---|---|
| | Treatment Group | | Control Group | | Regression Adjusted T-C Group Difference[b] | p-value to Test Difference | BH-Corrected Critical Value[c] | Estimated Effect Size |
| | Mean | Standard Deviation[a] | Mean | Standard Deviation | | | | |
| Delinquency/Misconduct (Range 1–4) | 3.41 | 0.74 | 3.40 | 0.74 | 0.00 | 0.84 | 0.05 | 0.01 |
| Gang Activity—Percent | 5.85 | | 5.46 | | 0.49 | 0.59 | 0.04 | 1.09 [d] |
| Tobacco Use—Percent | 6.11 | | 7.51 | | -1.17 | 0.24 | 0.02 | 0.83 [d] |
| Alcohol Use—Percent | 12.04 | | 13.38 | | -1.45 | 0.27 | 0.02 | 0.88 [d] |
| Drug Use—Percent | 7.44 | | 8.09 | | -0.66 | 0.53 | 0.04 | 0.91 [d] |
| Number of students | 1163 | | 1197 | | | | | |
| Percent missing data | ≤1% | | ≤1% | | | | | |
| **School-Reported Behavioral Outcomes** | | | | | | | | |
| *Truancy—Unexcused Absence Rate (Percent)* [e] | 2.04 | 4.80 | 2.47 | 6.91 | -0.45* | 0.02 | 0.01 | -0.14 |
| Number of students | 1163 | | 1197 | | | | | |
| Percent missing data | 42% | | 41% | | | | | |
| *Misconduct* [f] | | | | | | | | |
| Percent committing any infraction | 25.00 | | 22.91 | | 2.56 | 0.13 | 0.01 | 1.15 [d] |
| Percent committing repeated infractions (2+) | 14.21 | | 15.63 | | -0.98 | 0.48 | 0.03 | 0.93 [d] |
| *Delinquency* [f] | | | | | | | | |
| Percent committing any infraction | 18.13 | | 20.03 | | -1.51 | 0.35 | 0.03 | 0.91 [d] |
| Percent committing repeated infractions (2+) | 8.64 | | 9.13 | | -0.56 | 0.65 | 0.05 | 0.93 [d] |
| Number of students | 1163 | | 1197 | | | | | |
| Percent missing data | ≤22% | | ≤23% | | | | | |

[a]   Standard Deviations are only reported for Means or Mean Percents.

[b]   Regression Adjusted T-C Difference will not necessarily be equal to the difference between the Unadjusted Mean Outcomes.

[c]   Based on Benjamini-Hochberg test.

[d]   Odds ratio.

[e]   Based on 27 sites that reported unexcused absences and total days enrolled

[f]   Lower percents of the school-reported Truancy, Misconduct, and Delinquency items indicate more positive outcomes.

Treatment Group: Missing data ≤38%; Control Group:  Missing data ≤36%

*    p-value (of adjusted difference in means) < 0.05, two-tailed test.

+    p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit D.5: Subgroup Findings by Gender: Interpersonal Relationships, Personal Responsibility, and Community Involvement**

| Self-Reported Outcome | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | | |
| | Boys | | | | Girls | | | | | | | | |
| | Treatment | | Control | | Treatment | | Control | | Estimated Impact on Boys[b] | Estimated Impact on Girls[b] | Difference in Impacts | p-value to Test Difference | BH-Corrected Critical Value[c] |
| | Mean | Standard Deviation[a] | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Peer Relationships | 3.01 | 1.29 | 3.02 | 1.34 | 2.95 | 1.33 | 2.93 | 1.36 | 0.04 | 0.02 | 0.02 | 0.78 | 0.04 |
| Parental Relationships | 2.83 | 1.16 | 2.87 | 1.14 | 3.06 | 1.08 | 3.00 | 1.08 | -0.05 | 0.05 | -0.10 | 0.07 | 0.03 |
| Relationships with Other Adults | 2.22 | 1.21 | 2.15 | 1.17 | 2.33 | 1.16 | 2.21 | 1.13 | 0.07 | 0.08 | 0.00 | 0.95 | 0.05 |
| Personal Initiative | 2.66 | 0.97 | 2.74 | 0.89 | 2.81 | 0.94 | 2.76 | 0.93 | -0.06 | 0.04 | -0.10* | 0.03 | 0.01 |
| Number of students | 542 | 573 | 621 | 624 | | | | | | | | | |
| Percent missing data | ≤3% | ≤2% | ≤3% | ≤3% | | | | | | | | | |

[a]    Standard Deviations are only reported for Means or Mean Percents.

[b]    Estimated Impacts on Boys and Girls will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

[c]    Based on Benjamini-Hochberg test.

*    p-value (of adjusted difference in means) < 0.05, two-tailed test.
+    p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007.

**Exhibit D.6: Subgroup Findings by Gender: Academic Outcomes**

| | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | | |
| | Boys | | | | Girls | | | | | | | | |
| | Treatment | | Control | | Treatment | | Control | | | | | | BH-Corrected |
| Self-Reported Outcome (Range 1–4) | Mean | Standard Deviation[a] | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | Estimated Impact on Boys[b] | Estimated Impact on Girls[b] | Difference in Impacts | p-value to Test Difference | Critical Value[c] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Scholastic Efficacy | 2.99 | 0.74 | 2.99 | 0.83 | 3.05 | 0.77 | 2.95 | 0.78 | -0.01 | 0.10*+ | -0.11*+ | 0.01 | 0.01 |
| School Bonding | 2.85 | 0.97 | 2.90 | 1.02 | 3.05 | 0.89 | 2.97 | 0.98 | -0.05 | 0.09*+ | -0.14*+ | 0.01 | 0.01 |
| Future Orientation | 3.83 | 0.52 | 3.75 | 0.72 | 3.86 | 0.56 | 3.84 | 0.53 | 0.07*+ | 0.00 | 0.07* | 0.03 | 0.02 |
| Number of students | 542 | | 573 | | 621 | | 624 | | | | | | |
| Percent missing data | ≤2% | | ≤2% | | ≤3% | | ≤3% | | | | | | |
| **School-Reported Outcome** | | | | | | | | | | | | | |
| *Overall Absenteeism Rate (Percent)* d | 5.06 | 7.35 | 5.08 | 8.24 | 5.01 | 8.02 | 5.86 | 10.75 | -0.12 | -0.82* | 0.69 | 0.13 | 0.03 |
| Number of students | 542 | | 573 | | | | 624 | | | | | | |
| Percent missing data | 13% | | 16% | | 621 | | 19% | | | | | | |
| *Grades (Range 1–5)* e | | | | | 16% | | | | | | | | |
| Math | 3.16 | 1.79 | 3.19 | 1.71 | 3.29 | 1.74 | 3.34 | 1.78 | -0.07 | -0.04 | -0.03 | 0.74 | 0.05 |
| English Language Arts | 3.54 | 1.82 | 3.56 | 1.78 | 3.83 | 2.17 | 3.92 | 2.00 | -0.08 | 0.05 | -0.13 | 0.16 | 0.03 |
| Science | 3.47 | 2.05 | 3.49 | 2.11 | 3.78 | 2.12 | 3.83 | 2.07 | -0.04 | -0.03 | -0.01 | 0.88 | 0.05 |
| Social Studies | 3.43 | 1.97 | 3.37 | 1.98 | 3.77 | 2.11 | 3.91 | 2.01 | 0.01 | -0.05 | 0.06 | 0.53 | 0.04 |
| Number of students | 542 | | 573 | | 621 | | 624 | | | | | | |
| Percent missing data | ≤39% | | ≤34% | | ≤33% | | ≤35% | | | | | | |
| **State Assessment Tests** | | | | | | | | | | | | | |
| Math—Percent Proficient | 47.44 | | 47.61 | | 44.09 | | 46.60 | | -2.38 | -1.17 | -1.21 | 0.75 | 0.04 |
| Reading/ELA—Percent Proficient | 50.09 | | 48.57 | | 48.78 | | 52.87 | | 1.61 | -4.91 | 6.52 | 0.08 | 0.02 |
| Number of students | 542 | | 573 | | 621 | | 624 | | | | | | |
| Percent missing data | ≤23% | | ≤18% | | ≤25% | | ≤22% | | | | | | |

a    Standard Deviations are only reported for Means or Mean Percents.

b    Estimated Impacts on Boys and Girls will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

c    Based on Benjamini-Hochberg test.

d    Lower Overall Absenteeism Rates indicate more positive outcomes.

e    Higher scores indicate higher grades; see Appendix F for further explanation of how these scores were derived.

\*    p-value (of adjusted difference in means) < 0.05, two-tailed test.

+    p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing..

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit D.7: Subgroup Findings by Gender: Delinquent Behaviors and Participation in Harmful Activities**

| Self-Reported Outcome[a] | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Boys | | | | Girls | | | | | | | | |
| | Treatment | | Control | | Treatment | | Control | | Estimated Impact on Boys[c] | Estimated Impact on Girls[c] | Difference in Impacts | p-value to Test Difference | BH-Corrected Critical Value[d] |
| | Mean | Standard Deviation[b] | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | | | | | |
| Delinquency/Misconduct (Range 1–4) | 3.37 | 0.74 | 3.33 | 0.82 | 3.45 | 0.73 | 3.47 | 0.65 | 0.04 | -0.02 | 0.06 | 0.11 | 0.02 |
| Gang Activity—Percent | 9.38 | | 8.51 | | 2.72 | | 2.66 | | 1.18 | -0.09 | 1.27 | 0.50 | 0.04 |
| Tobacco Use—Percent | 6.86 | | 8.15 | | 5.44 | | 6.93 | | -0.76 | -2.23 | 1.48 | 0.47 | 0.04 |
| Alcohol Use—Percent | 11.99 | | 13.50 | | 12.09 | | 13.27 | | -1.09 | -1.93 | 0.83 | 0.75 | 0.05 |
| Drug Use—Percent | 8.65 | | 8.33 | | 6.36 | | 7.88 | | 0.93 | -2.55 | 3.48 | 0.10 | 0.01 |
| Number of students | 542 | | 573 | | 621 | | 624 | | | | | | |
| Percent missing data | ≤4% | | ≤5% | | ≤3% | | ≤3% | | | | | | |
| **School-Reported Behavioral Outcome** | | | | | | | | | | | | | |
| *Truancy—Unexcused Absence Rate (Percent)* [e,f] | 2.03 | 4.35 | 2.13 | 5.76 | 2.06 | 5.24 | 2.85 | 7.97 | -0.23 | -0.76* | 0.53 | 0.20 | 0.02 |
| Number of students | 542 | | 573 | | 621 | | 624 | | | | | | |
| Percent missing data | 37% | | 36% | | 47% | | 47% | | | | | | |
| *Misconduct* [f] | | | | | | | | | | | | | |
| Percent committing any infraction | 31.59 | | 25.64 | | 18.87 | | 20.36 | | 6.33* | -1.21 | 7.54* | 0.03 | 0.01 |
| Percent committing repeated infractions (2+) | 18.13 | | 17.63 | | 10.56 | | 13.76 | | 0.83 | -2.55 | 3.38 | 0.23 | 0.03 |
| *Delinquency* [f] | | | | | | | | | | | | | |
| Percent committing any infraction | 22.73 | | 24.02 | | 13.85 | | 16.30 | | 0.16 | -2.55 | 2.71 | 0.41 | 0.03 |
| Percent committing repeated infractions (2+) | 10.18 | | 10.58 | | 7.21 | | 7.77 | | 0.61 | -0.97 | 1.59 | 0.52 | 0.05 |
| Number of students | 542 | | 573 | | 621 | | 624 | | | | | | |
| Percent missing data | ≤19% | | ≤21% | | ≤23% | | ≤23% | | | | | | |

[a]  Higher scores on the Misconduct and Delinquency scales indicate more positive outcomes.

[b]  Standard Deviations are only reported for Means or Mean Percents.

[c]  Estimated Impacts on Boys and Girls will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

[d]  Based on Benjamini-Hochberg test.

[e]  Based on 27 sites that reported unexcused absences and total days enrolled.

[f]  Lower percents of the school-reported Truancy, Misconduct, and  Delinquency items indicate more positive outcomes.

\*  p-value (of adjusted difference in means) < 0.05, two-tailed test.

+  p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➜ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit D.8: Subgroup Findings by Age: Interpersonal Relationships, Personal Responsibility, and Community Involvement**

| | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | | |
| | Students Below Age 12 | | | | Students Aged 12 and Older | | | | | | | | |
| | Treatment | | Control | | Treatment | | Control | | | | | | |
| Self-Reported Outcome | Mean | Standard Deviation[a] | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | Estimated Impact on Younger Students[b] | Estimated Impact on Older Students[b] | Difference in Impacts | p-value to Test Difference | BH-Corrected Critical Value[c] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Peer Relationships | 2.88 | 1.34 | 2.88 | 1.37 | 3.21 | 1.16 | 3.19 | 1.25 | 0.02 | 0.05 | -0.03 | 0.61 | 0.03 |
| Parental Relationships | 3.02 | 1.13 | 3.03 | 1.10 | 2.79 | 1.10 | 2.71 | 1.09 | -0.03 | 0.10* | -0.13* | 0.03 | 0.01 |
| Relationships with Other Adults | 2.22 | 1.18 | 2.14 | 1.19 | 2.42 | 1.17 | 2.29 | 1.04 | 0.07 | 0.09 | -0.02 | 0.76 | 0.04 |
| Personal Initiative | 2.78 | 0.96 | 2.80 | 0.91 | 2.63 | 0.94 | 2.64 | 0.89 | -0.01 | -0.02 | 0.01 | 0.87 | 0.05 |
| Number of students | 826 | | 833 | | 337 | | 364 | | | | | | |
| Percent missing data | ≤3% | | ≤3% | | ≤2% | | ≤2% | | | | | | |

[a]  Standard Deviations are only reported for Means or Mean Percents.

[b]  Estimated Impacts on Younger Students and Older Students will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

[c]  Based on Benjamini-Hochberg test.

*  p-value (of adjusted difference in means) < 0.05, two-tailed test.

+  p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007.

### Exhibit D.9: Subgroup Findings by Age: Academic Outcomes

| Self-Reported Outcome (Range 1–4) | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | | |
| | Students Below Age 12 | | | | Students Aged 12 and Older | | | | | | | | |
| | Treatment | | Control | | | | | | | | | | |
| | Mean | Standard Deviation[a] | Mean | Standard Deviation[a] | Mean | Standard Deviation | Mean | Standard Deviation | Estimated Impact on Younger Students[b] | Estimated Impact on Older Students[b] | Difference in Impacts | p-value to Test Difference | BH-Corrected Critical Value[c] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Scholastic Efficacy | 3.06 | 0.74 | 3.00 | 0.78 | 2.93 | 0.79 | 2.90 | 0.85 | 0.05* | 0.03 | 0.03 | 0.56 | 0.02 |
| School Bonding | 3.04 | 0.90 | 3.04 | 0.96 | 2.76 | 0.95 | 2.70 | 1.01 | 0.00 | 0.08 | -0.08 | 0.11 | 0.01 |
| Future Orientation | 3.87 | 0.47 | 3.83 | 0.59 | 3.78 | 0.66 | 3.74 | 0.71 | 0.03 | 0.04 | -0.01 | 0.84 | 0.04 |
| Number of students | 826 | | 833 | | 337 | | 364 | | | | | | |
| Percent missing data | ≤3% | | ≤3% | | ≤2% | | ≤2% | | | | | | |
| **School-Reported Outcome** | | | | | | | | | | | | | |
| *Overall Absenteeism Rate (Percent)* [d] | 4.24 | 5.99 | 4.64 | 7.95 | 6.94 | 10.42 | 7.50 | 12.18 | -0.51* | -0.88 | 0.37 | 0.54 | 0.02 |
| Number of students | 826 | | 833 | | 337 | | 364 | | | | | | |
| Percent missing data | 14% | | 18% | | 17% | | 16% | | | | | | |
| *Grades (Range 1–5)* [e] | | | | | | | | | | | | | |
| Math | 3.37 | 1.71 | 3.43 | 1.65 | 2.91 | 1.86 | 2.89 | 1.75 | -0.07 | -0.03 | -0.04 | 0.71 | 0.03 |
| English Language Arts | 3.85 | 1.72 | 3.88 | 1.66 | 3.15 | 1.95 | 3.16 | 1.78 | -0.04 | -0.05 | 0.01 | 0.89 | 0.05 |
| Science | 3.93 | 1.91 | 4.00 | 1.84 | 3.11 | 2.18 | 2.99 | 1.94 | -0.02 | -0.02 | 0.00 | 0.98 | 0.05 |
| Social Studies | 3.89 | 1.99 | 3.90 | 1.91 | 3.11 | 2.03 | 3.08 | 1.72 | 0.02 | -0.01 | 0.03 | 0.78 | 0.03 |
| Number of students | 826 | | 833 | | 337 | | 364 | | | | | | |
| Percent missing data | ≤37% | | ≤35% | | ≤34% | | ≤32% | | | | | | |
| *State Assessment Tests* | | | | | | | | | | | | | |
| Math—Percent Proficient | 50.57 | | 52.84 | | 34.21 | | 33.21 | | -3.24 | 1.13 | -4.37 | 0.28 | 0.01 |
| Reading/ELA—Percent Proficient | 52.96 | | 55.94 | | 40.90 | | 38.39 | | -1.83 | -1.12 | -0.71 | 0.86 | 0.04 |
| Number of students | 826 | | 833 | | 337 | | 364 | | | | | | |
| Percent missing data | ≤24% | | ≤20% | | ≤26% | | ≤21% | | | | | | |

[a] Standard Deviations are only reported for Means or Mean Percents.

[b] Estimated Impacts on Younger Students and Older Students will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

[c] Based on Benjamini-Hochberg test.

[d] Lower Overall Absenteeism Rates indicate more positive outcomes.

[e] Higher scores indicate higher grades; see Appendix F for further explanation of how these scores were derived.

\* p-value (of adjusted difference in means) < 0.05, two-tailed test.

\+ p-value (of adjusted difference in means) < BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing..

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit D.10: Subgroup Findings by Age: Delinquent Behaviors and Participation in Harmful Activities**

| Self-Reported Outcome[a] | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Students Below Age 12 | | | | Students Aged 12 and Older | | | | Estimated Impact on Younger Students[b] | Estimated Impact on Older Students[b] | Difference in Impacts | p-value to Test Difference | BH-Corrected Critical Value[d] |
| | Treatment | | Control | | | | | | | | | | |
| | Mean | Standard Deviation[b] | Mean | Standard Deviation[b] | Mean | Standard Deviation | Mean | Standard Deviation | | | | | |
| Delinquency/Misconduct (Range 1–4) | 3.49 | 0.66 | 3.47 | 0.71 | 3.22 | 0.83 | 3.25 | 0.77 | 0.02 | -0.04 | 0.06 | 0.14 | 0.02 |
| Gang Activity—Percent | 5.56 | | 4.59 | | 6.54 | | 7.53 | | 0.89 | -0.92 | 1.81 | 0.41 | 0.03 |
| Tobacco Use—Percent | 4.08 | | 5.47 | | 10.79 | | 12.29 | | -1.73 | 0.12 | -1.84 | 0.48 | 0.04 |
| Alcohol Use—Percent | 8.81 | | 9.19 | | 19.46 | | 23.17 | | -0.21 | -3.51 | 3.31 | 0.32 | 0.03 |
| Drug Use—Percent | 4.77 | | 5.43 | | 13.66 | | 14.42 | | -0.70 | -1.30 | 0.60 | 0.82 | 0.05 |
| Number of students | 826 | | 833 | | 337 | | 364 | | | | | | |
| Percent missing data | ≤4% | | ≤4% | | ≤2% | | ≤3% | | | | | | |
| **School-Reported Behavioral Outcome** | | | | | | | | | | | | | |
| ***Truancy—Unexcused Absence Rate (Percent)*** [e,f] | 1.60 | 3.67 | 1.90 | 5.44 | 2.98 | 6.50 | 3.78 | 9.06 | -0.55*+ | -0.84 | 0.30 | 0.58 | 0.04 |
| Number of students | 826 | | 833 | | 337 | | 364 | | | | | | |
| Percent missing data | ≤43% | | ≤43% | | ≤39% | | ≤38% | | | | | | |
| ***Misconduct*** [f] | | | | | | | | | | | | | |
| Percent committing any infraction | 22.67 | | 19.49 | | 31.00 | | 31.61 | | 4.11* | -3.03 | 7.14 | 0.06 | 0.01 |
| Percent committing repeated infractions (2+) | 12.05 | | 12.63 | | 19.77 | | 23.27 | | 0.46 | -6.32* | 6.78* | 0.03 | 0.01 |
| ***Delinquency*** [f] | | | | | | | | | | | | | |
| Percent committing any infraction | 14.75 | | 15.25 | | 26.83 | | 32.18 | | -0.22 | -6.40 | 6.17 | 0.12 | 0.02 |
| Percent committing repeated infractions (2+) | 5.01 | | 5.46 | | 17.99 | | 18.47 | | -0.56 | -1.09 | 0.53 | 0.87 | 0.05 |
| Number of students | 826 | | 833 | | 337 | | 364 | | | | | | |
| Percent missing data | ≤20% | | ≤20% | | ≤24% | | ≤28% | | | | | | |

[a]  Higher scores on the Misconduct and Delinquency scales indicate more positive outcomes.

[b]  Standard Deviations are only reported for Means or Mean Percents.

[c]  Estimated Impacts on Younger Students and Older Students will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

[d]  Based on Benjamini-Hochberg test.

[e]  Based on 27 sites that reported unexcused absences and total days enrolled.

[f]  Lower percents of the school-reported Truancy, Misconduct, and Delinquency items indicate more positive outcomes.

\*  p-value (of adjusted difference in means) < 0.05, two-tailed test.

+  p-value (of adjusted difference in means) < BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit D.11: Subgroup Findings by Family Structure: Interpersonal Relationships, Personal Responsibility, and Community Involvement**

| | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | | |
| | Two-Parent Households | | | | Other Households | | | | | | | | |
| | Treatment | | Control | | Treatment | | Control | | | | | | |
| Self-Reported Outcome | Mean | Standard Deviation[a] | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | Estimated Impact on Two-Parent HHs[c] | Estimated Impact on Other HHs[b] | Difference in Impacts | p-value to Test Difference | BH-Corrected Critical Value[c] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Peer Relationships | 2.93 | 1.37 | 2.88 | 1.38 | 3.04 | 1.22 | 3.10 | 1.29 | 0.09 | -0.03 | 0.11 | 0.10 | 0.03 |
| Parental Relationships | 3.03 | 1.09 | 2.97 | 1.12 | 2.85 | 1.17 | 2.89 | 1.10 | 0.05 | -0.06 | 0.11 | 0.06 | 0.01 |
| Relationships with Other Adults | 2.25 | 1.20 | 2.15 | 1.13 | 2.31 | 1.17 | 2.22 | 1.18 | 0.07 | 0.07 | 0.01 | 0.90 | 0.05 |
| Personal Initiative | 2.77 | 0.95 | 2.75 | 0.90 | 2.70 | 0.96 | 2.76 | 0.92 | 0.02 | -0.06 | 0.08 | 0.13 | 0.04 |
| Number of students | 705 | | 735 | | 549 | | 556 | | | | | | |
| Percent missing data | ≤9% | | ≤8% | | ≤11% | | ≤12% | | | | | | |

[a] Standard Deviations are only reported for Means or Mean Percents.

[b] Estimated Impacts on Two-Parent Households and Other Households will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

[c] Based on Benjamini-Hochberg test.

* p-value (of adjusted difference in means) < 0.05, two-tailed test.

+ p-value (of adjusted difference in means) < BH-Corrected Critical Value ➜ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007.

**Exhibit D.12: Subgroup Findings by Family Structure: Academic Outcomes**

| | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | | |
| | Two-Parent Households | | | | Other Households | | | | | | | | |
| | Treatment | | Control | | Treatment | | Control | | Estimated Impact on Two-Parent HHs[b] | Estimated Impact on Other HHs[b] | Difference in Impacts | p-value to Test Difference | BH-Corrected Critical Value[c] |
| Self-Reported Outcome (Range 1–4) | Mean | Standard Deviation[a] | Mean | Standard Deviation[a] | Mean | Standard Deviation | Mean | Standard Deviation | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Scholastic Efficacy | 3.02 | 0.78 | 2.96 | 0.82 | 3.02 | 0.73 | 2.98 | 0.79 | 0.05 | 0.05 | 0.00 | 0.91 | 0.05 |
| School Bonding | 2.99 | 0.95 | 2.97 | 0.97 | 2.91 | 0.91 | 2.89 | 1.03 | 0.00 | 0.03 | -0.03 | 0.60 | 0.02 |
| Future Orientation | 3.85 | 0.56 | 3.80 | 0.62 | 3.83 | 0.52 | 3.81 | 0.65 | 0.03 | 0.03 | 0.01 | 0.87 | 0.04 |
| Number of students | 705 | | 735 | | 549 | | 556 | | | | | | |
| Percent missing data | ≤9% | | ≤8% | | ≤11% | | ≤13% | | | | | | |
| **School-Reported Outcome** | | | | | | | | | | | | | |
| *Overall Absenteeism Rate (Percent)* [d] | 4.37 | 6.34 | 4.58 | 8.01 | 5.75 | 8.83 | 6.52 | 10.93 | -0.19 | -0.65* | 0.46 | 0.33 | 0.01 |
| Number of students | 705 | | 735 | | 549 | | 556 | | | | | | |
| Percent missing data | 22% | | 24% | | 21% | | 23% | | | | | | |
| *Grades (Range 1–5)* [e] | | | | | | | | | | | | | |
| Math | 3.32 | 1.86 | 3.33 | 1.82 | 3.30 | 1.76 | 3.25 | 1.76 | -0.09 | -0.01 | -0.09 | 0.34 | 0.01 |
| English Language Arts | 3.71 | 1.91 | 3.70 | 1.85 | 3.61 | 1.92 | 3.60 | 1.67 | -0.03 | -0.02 | -0.01 | 0.88 | 0.04 |
| Science | 3.80 | 2.33 | 3.72 | 2.29 | 3.67 | 2.25 | 3.61 | 1.83 | -0.04 | -0.05 | 0.01 | 0.92 | 0.05 |
| Social Studies | 3.82 | 2.28 | 3.71 | 2.23 | 3.55 | 2.22 | 3.62 | 2.06 | 0.04 | 0.02 | 0.02 | 0.83 | 0.03 |
| Number of students | 705 | | 735 | | 549 | | 556 | | | | | | |
| Percent missing data | ≤42% | | ≤41% | | ≤41% | | ≤37% | | | | | | |
| **State Assessment Tests** | | | | | | | | | | | | | |
| Math—Percent Proficient | 46.70 | | 50.69 | | 44.29 | | 42.22 | | -2.38 | -1.17 | -1.21 | 0.75 | 0.03 |
| Reading/ELA—Percent Proficient | 48.81 | | 50.39 | | 50.31 | | 51.12 | | 1.61 | -4.91 | 6.52 | 0. 08 | 0.01 |
| Number of students | 705 | | 735 | | 549 | | 556 | | | | | | |
| Percent missing data | ≤41% | | ≤36% | | ≤15% | | ≤12% | | | | | | |

a   Standard Deviations are only reported for Means or Mean Percents.

b   Estimated Impacts on Two-Parent Households and Other Households will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

c   Based on Benjamini-Hochberg test.

d   Lower Overall Absenteeism Rates indicate more positive outcomes.

e   Higher scores indicate higher grades; see Appendix F for further explanation of how these scores were derived.

*   p-value (of adjusted difference in means) < 0.05, two-tailed test.

+   p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing..

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit D.13: Subgroup Findings by Family Structure: Delinquent Behaviors and Participation in Harmful Activities**

| | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | | |
| | Two-Parent Households | | | | Other Households | | | | | | | | |
| | Treatment | | Control | | Treatment | | Control | | Estimated Impact on Two-Parent HHs[c] | Estimated Impact on Other HHs[c] | Difference in Impacts | p-value to Test Difference | BH-Corrected Critical Value[d] |
| Self-Reported Outcome[a] | Mean | Standard Deviation[b] | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Delinquency/Misconduct (Range 1–4) | 3.45 | 0.74 | 3.43 | 0.75 | 3.36 | 0.72 | 3.37 | 0.73 | 0.00 | 0.00 | -0.01 | 0.89 | 0.05 |
| Gang Activity—Percent | 5.77 | | 5.60 | | 5.96 | | 5.26 | | 0.12 | 1.07 | -0.96 | 0.61 | 0.03 |
| Tobacco Use—Percent | 5.41 | | 7.11 | | 7.02 | | 8.05 | | -1.24 | -1.15 | -0.10 | 0.96 | 0.05 |
| Alcohol Use—Percent | 10.24 | | 14.35 | | 14.39 | | 12.06 | | -3.40* | 1.20 | -4.60 | 0.09 | 0.01 |
| Drug Use—Percent | 7.25 | | 7.96 | | 7.68 | | 8.27 | | -0.14 | -0.87 | 0.73 | 0.74 | 0.04 |
| Number of students | 705 | 735 | 549 | 556 | | | | | | | | | |
| Percent missing data | ≤10% | ≤9% | ≤11% | ≤13% | | | | | | | | | |
| **School-Reported Behavioral Outcome** | | | | | | | | | | | | | |
| *Truancy—Unexcused Absence Rate (Percent)* [e,f] | 1.83 | 4.07 | 1.72 | 4.91 | 2.33 | 5.59 | 3.38 | 8.59 | 0.02 | 0.91* | -0.90* | 0.03 | 0.01 |
| Number of students | 705 | 735 | 549 | 556 | | | | | | | | | |
| Percent missing data | 47% | 48% | 46% | 44% | | | | | | | | | |
| *Misconduct* [f] | | | | | | | | | | | | | |
| Percent committing any infraction | 21.41 | | 19.67 | | 29.42 | | 27.11 | | 1.76 | 2.83 | -1.07 | 0.76 | 0.04 |
| Percent committing repeated infractions (2+) | 10.68 | | 13.78 | | 18.31 | | 17.98 | | -2.16 | 0.30 | -2.46 | 0.39 | 0.02 |
| *Delinquency* [f] | | | | | | | | | | | | | |
| Percent committing any infraction | 13.97 | | 17.60 | | 23.06 | | 23.40 | | -3.56 | -0.06 | -3.50 | 0.30 | 0.02 |
| Percent committing repeated infractions (2+) | 6.17 | | 7.60 | | 11.57 | | 11.18 | | -1.92 | 0.15 | -2.08 | 0.41 | 0.03 |
| Number of students | 705 | | 735 | | 549 | | 556 | | | | | | |
| Percent missing data | ≤30% | | ≤31% | | ≤25% | | ≤25% | | | | | | |

[a] Higher scores on the Misconduct and Delinquency scales indicate more positive outcomes.

[b] Standard Deviations are only reported for Means or Mean Percents.

[c] Estimated Impacts on Two-Parent Households and Other Households will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

[d] Based on Benjamini-Hochberg test.

[e] Based on 27 sites that reported unexcused absences and total days enrolled.

[f] Lower percents of the school-reported Truancy, Misconduct, and Delinquency items indicate more positive outcomes.

\* p-value (of adjusted difference in means) < 0.05, two-tailed test.
\+ p-value (of adjusted difference in means) < BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit D.14: Subgroup Findings by Academic Risk: Interpersonal Relationships, Personal Responsibility, and Community Involvement**

| Self-Reported Outcome | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | | |
| | Proficient Students | | | | Not Proficient Students | | | | | | | | |
| | Treatment | | Control | | Treatment | | Control | | Estimated Impact on Proficient Students[b] | Estimated Impact on Not Proficient Students[b] | Difference in Impacts | p-value to Test Difference | BH-Corrected Critical Value[c] |
| | Mean | Standard Deviation[a] | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | | | | | |
| Peer Relationships | 3.04 | 1.25 | 3.08 | 1.26 | 2.95 | 1.36 | 2.91 | 1.39 | 0.05 | -0.01 | 0.05 | 0.47 | 0.03 |
| Parental Relationships | 3.01 | 1.13 | 2.92 | 1.12 | 2.91 | 1.11 | 2.94 | 1.12 | 0.03 | 0.00 | 0.03 | 0.60 | 0.04 |
| Relationships with Other Adults | 2.17 | 1.16 | 2.11 | 1.12 | 2.36 | 1.18 | 2.23 | 1.18 | 0.01 | 0.13*+ | -0.12 | 0.09 | 0.01 |
| Personal Initiative | 2.76 | 0.98 | 2.78 | 0.89 | 2.72 | 0.95 | 2.74 | 0.92 | -0.04 | -0.02 | -0.02 | 0.71 | 0.05 |
| Number of students | 373 | | 398 | | 550 | | 586 | | | | | | |
| Percent missing data | ≤2% | | ≤2% | | ≤3% | | ≤2% | | | | | | |

[a]   Standard Deviations are only reported for Means or Mean Percents.

[b]   Estimated Impacts on Proficient and Not Proficient students will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

[c]   Based on Benjamini-Hochberg test.

*    p-value (of adjusted difference in means) < 0.05, two-tailed test.
+    p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007.

**Exhibit D.15: Subgroup Findings by Academic Risk: Academic Outcomes**

| | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | | |
| | Proficient Students | | | | Not Proficient Students | | | | | | | | |
| | Treatment | | Control | | Treatment | | Control | | | | | | |
| Self-Reported Outcome (Range 1–4) | Mean | Standard Deviation[a] | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | Estimated Impact on Proficient Students[b] | Estimated Impact on Not Proficient Students[b] | Difference in Impacts | p-value to Test Difference | BH-Corrected Critical Value[c] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Scholastic Efficacy | 3.11 | 0.76 | 3.07 | 0.78 | 2.95 | 0.73 | 2.90 | 0.81 | 0.05 | 0.05 | 0.00 | 0.98 | 0.05 |
| School Bonding | 2.99 | 0.96 | 2.95 | 0.98 | 2.93 | 0.91 | 2.92 | 1.01 | -0.01 | 0.03 | -0.05 | 0.39 | 0.02 |
| Future Orientation | 3.88 | 0.44 | 3.83 | 0.60 | 3.83 | 0.58 | 3.79 | 0.67 | 0.03 | 0.04 | -0.01 | 0.73 | 0.03 |
| Number of students | 373 | | 398 | | 550 | | 586 | | | | | | |
| Percent missing data | ≤2% | | ≤2% | | ≤3% | | ≤3% | | | | | | |
| **School-Reported Outcome** | | | | | | | | | | | | | |
| *Overall Absenteeism Rate (Percent)* [d] | 4.16 | 6.20 | 4.91 | 8.50 | 5.19 | 7.55 | 5.81 | 9.81 | -0.60 | -0.72* | 0.13 | 0.79 | 0.04 |
| Number of students | 373 | | 398 | | 550 | | 586 | | | | | | |
| Percent missing data | 6% | | 10% | | 10% | | 14% | | | | | | |
| *Grades (Range 1–5)* [e] | | | | | | | | | | | | | |
| Math | 4.04 | 1.93 | 4.12 | 1.99 | 3.18 | 1.84 | 3.12 | 1.72 | -0.03 | -0.02 | 0.00 | 0.99 | 0.05 |
| English Language Arts | 4.53 | 2.60 | 4.57 | 2.52 | 3.61 | 1.82 | 3.63 | 1.72 | 0.01 | -0.07 | 0.09 | 0.39 | 0.01 |
| Science | 4.46 | 2.45 | 4.61 | 2.54 | 3.62 | 2.21 | 3.60 | 2.12 | -0.08 | -0.04 | -0.05 | 0.66 | 0.03 |
| Social Studies | 4.48 | 2.81 | 4.61 | 2.69 | 3.43 | 1.97 | 3.37 | 1.98 | 0.07 | 0.00 | 0.07 | 0.51 | 0.02 |
| Number of students | 373 | | 398 | | 550 | | 586 | | | | | | |
| Percent missing data | ≤34% | | ≤32% | | ≤32% | | ≤31% | | | | | | |
| **State Assessment Tests** | | | | | | | | | | | | | |
| Math—Percent Proficient | 75.04 | | 78.58 | | 26.19 | | 25.36 | | -3.35 | 0.33 | -3.68 | 0.33 | 0.01 |
| Reading/ELA—Percent Proficient | 80.50 | | 83.40 | | 28.83 | | 28.46 | | -1.75 | -0.53 | -1.23 | 0.74 | 0.04 |
| Number of students | 373 | | 398 | | 550 | | 586 | | | | | | |
| Percent missing data | ≤5% | | ≤2% | | ≤4% | | ≤4% | | | | | | |

[a] Standard Deviations are only reported for Means or Mean Percents.

[b] Estimated Impacts on Proficient and Not Proficient students will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

[c] Based on Benjamini-Hochberg test.

[d] Lower Overall Absenteeism Rates indicate more positive outcomes.

[e] Higher scores indicate higher grades; see Appendix F for further explanation of how these scores were derived.

\* p-value (of adjusted difference in means) < 0.05, two-tailed test.

+ p-value (of adjusted difference in means) < BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing..

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit D.16: Subgroup Findings by Academic Risk: Delinquent Behaviors and Participation in Harmful Activities**

| Self-Reported Outcome[a] | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Proficient Students | | | | Not Proficient Students | | | | Estimated Impact on Proficient Students[c] | Estimated Impact on Not Proficient Students[c] | Difference in Impacts | p-value to Test Difference | BH-Corrected Critical Value[d] |
| | Treatment | | Control | | Treatment | | Control | | | | | | |
| | Mean | Standard Deviation[b] | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | | | | | |
| Delinquency/Misconduct (Range 1–4) | 3.45 | 0.70 | 3.43 | 0.73 | 3.38 | 0.75 | 3.39 | 0.76 | 0.02 | -0.01 | 0.03 | 0.49 | 0.01 |
| Gang Activity—Percent | 4.26 | | 2.97 | | 6.89 | | 6.21 | | 1.24 | 1.31 | -0.07 | 0.97 | 0.05 |
| Tobacco Use—Percent | 4.34 | | 7.12 | | 6.57 | | 7.81 | | -1.55 | -1.29 | -0.27 | 0.91 | 0.05 |
| Alcohol Use—Percent | 11.10 | | 14.60 | | 11.69 | | 13.34 | | -3.29 | -1.25 | -2.03 | 0.50 | 0.02 |
| Drug Use—Percent | 6.67 | | 9.95 | | 6.59 | | 7.58 | | -1.72 | -1.26 | -0.46 | 0.85 | 0.04 |
| Number of students | 373 | | 398 | | 550 | | 586 | | | | | | |
| Percent missing data | ≤2% | | ≤2% | | ≤4% | | ≤6% | | | | | | |
| **School-Reported Behavioral Outcome** | | | | | | | | | | | | | |
| ***Truancy—Unexcused Absence Rate (Percent)*** [e,f] | 1.24 | 3.53 | 1.72 | 4.88 | 2.42 | 4.92 | 2.98 | 8.13 | -0.38 | -0.65* | 0.26 | 0.51 | 0.02 |
| Number of students | 373 | | 398 | | 550 | | 586 | | | | | | |
| Percent missing data | 39% | | 40% | | 36% | | 38% | | | | | | |
| ***Misconduct*** [f] | | | | | | | | | | | | | |
| Percent committing any infraction | 21.02 | | 19.70 | | 25.94 | | 25.77 | | 0.15 | 1.76 | -1.61 | 0.67 | 0.03 |
| Percent committing repeated infractions (2+) | 12.35 | | 13.76 | | 15.11 | | 17.93 | | -2.13 | -1.67 | -0.46 | 0.89 | 0.04 |
| ***Delinquency*** [f] | | | | | | | | | | | | | |
| Percent committing any infraction | 17.15 | | 17.86 | | 18.41 | | 22.48 | | 1.77 | -4.24 | 6.00 | 0.11 | 0.01 |
| Percent committing repeated infractions (2+) | 6.51 | | 8.84 | | 8.85 | | 10.36 | | -2.31 | -1.75 | -0.57 | 0.84 | 0.03 |
| Number of students | 373 | | 398 | | 550 | | 586 | | | | | | |
| Percent missing data | ≤27% | | ≤30% | | ≤13% | | ≤15% | | | | | | |

[a]  Higher scores on the Misconduct and Delinquency scales indicate more positive outcomes.

[b]  Standard Deviations are only reported for Means or Mean Percents.

[c]  Estimated Impacts on Proficient and Not Proficient students will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

[d]  Based on Benjamini-Hochberg test.

[e]  Based on 27 sites that reported unexcused absences and total days enrolled.

[f]  Lower percents of the school-reported Truancy, Misconduct, and  Delinquency items indicate more positive outcomes.

\*  p-value (of adjusted difference in means) < 0.05, two-tailed test.

+  p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit D.17: Subgroup Findings by Baseline Delinquency: Interpersonal Relationships, Personal Responsibility, and Community Involvement**

| | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | | |
| | Any Delinquency | | | | No Delinquency | | | | | | | | |
| | Treatment | | Control | | Treatment | | Control | | | | | | |
| Self-Reported Outcome | Mean | Standard Deviation[a] | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | Estimated Impact on Students Reporting Any Delinquency[b] | Estimated Impact on Students Reporting No Delinquency[a] | Difference in Impacts | p-value to Test Difference | BH-Corrected Critical Value[c] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Peer Relationships | 3.06 | 1.31 | 2.97 | 1.41 | 2.95 | 1.31 | 2.97 | | 0.02 | 0.00 | 0.03 | 0.59 | 0.03 |
| Parental Relationships | 2.69 | 1.14 | 2.67 | 1.20 | 3.03 | 1.10 | 3.03 | | -0.01 | 0.01 | -0.02 | 0.78 | 0.03 |
| Relationships with Other Adults | 2.39 | 1.16 | 2.27 | 1.15 | 2.24 | 1.19 | 2.15 | | 0.07 | 0.07* | -0.01 | 0.94 | 0.04 |
| Personal Initiative | 2.63 | 0.94 | 2.61 | 0.96 | 2.77 | 0.96 | 2.80 | | -0.01 | -0.01 | 0.00 | 0.95 | 0.05 |
| Number of students | 277 | | 310 | | 886 | | 887 | | | | | | |
| Percent missing data | ≤4% | | ≤2% | | ≤3% | | ≤2% | | | | | | |

[a]  Standard Deviations are only reported for Means or Mean Percents.

[b]  Estimated Impacts on Any Delinquency and No Delinquency students will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

[c]  Based on Benjamini-Hochberg test.

*  p-value (of adjusted difference in means) < 0.05, two-tailed test.
+  p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007.

**Exhibit D.18: Subgroup Findings by Baseline Delinquency: Academic Outcomes**

| | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | | |
| | Any Delinquency | | | | No Delinquency | | | | | | | | |
| | Treatment | | Control | | Treatment | | Control | | Estimated Impact on Students Reporting Any Delinquency[b] | Estimated Impact on Students Reporting No Delinquency[b] | Difference in Impacts | p-value to Test Difference | BH-Corrected Critical Value[c] |
| Self-Reported Outcome (Range 1–4) | Mean | Standard Deviation[a=] | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Scholastic Efficacy | 2.95 | 0.71 | 2.84 | 0.83 | 3.04 | 0.77 | 3.02 | 0.78 | 0.12*+ | 0.02 | 0.09* | 0.04 | 0.01 |
| School Bonding | 2.77 | 0.91 | 2.74 | 1.03 | 3.02 | 0.93 | 3.01 | 0.97 | 0.03 | 0.02 | 0.00 | 0.94 | 0.05 |
| Future Orientation | 3.76 | 0.69 | 3.67 | 0.81 | 3.87 | 0.48 | 3.84 | 0.55 | 0.05 | 0.02 | 0.03 | 0.46 | 0.04 |
| Number of students | 277 | | 310 | | 886 | | 887 | | | | | | |
| Percent missing data | ≤3% | | ≤4% | | ≤3% | | ≤3% | | | | | | |
| **School-Reported Outcome** | | | | | | | | | | | | | |
| *Overall Absenteeism Rate (Percent)* [d] | 6.44 | 9.47 | 6.00 | 11.37 | 4.56 | 6.92 | 5.30 | 8.87 | -0.13 | -0.61* | 0.49 | 0.43 | 0.03 |
| Number of students | 277 | | 310 | | 886 | | 887 | | | | | | |
| Percent missing data | 12% | | 12% | | 16% | | 19% | | | | | | |
| *Grades (Range 1–5)* [e] | | | | | | | | | | | | | |
| Math | 3.09 | 1.74 | 3.06 | 1.73 | 3.24 | 1.70 | 3.31 | 1.67 | 0.00 | -0.08 | 0.07 | 0.49 | 0.04 |
| English Language Arts | 3.54 | 2.42 | 3.50 | 2.10 | 3.76 | 1.83 | 3.81 | 1.74 | 0.06 | -0.08 | 0.14 | 0.19 | 0.02 |
| Science | 3.32 | 2.11 | 3.29 | 1.93 | 3.73 | 1.96 | 3.74 | 1.97 | 0.03 | -0.06 | 0.09 | 0.44 | 0.03 |
| Social Studies | 3.30 | 1.99 | 3.25 | 1.86 | 3.74 | 2.02 | 3.79 | 1.93 | 0.09 | -0.07 | 0.16 | 0.16 | 0.01 |
| Number of students | 277 | | 310 | | 886 | | 887 | | | | | | |
| Percent missing data | ≤42% | | ≤36% | | ≤34% | | ≤34% | | | | | | |
| *State Assessment Tests* | | | | | | | | | | | | | |
| Math—Percent Proficient | 46.19 | | 43.20 | | 45.55 | | 48.40 | | 2.38 | -3.17 | 5.55 | 0.21 | 0.02 |
| Reading/ELA—Percent Proficient | 51.83 | | 48.46 | | 48.68 | | 51.52 | | -0.03 | -2.19 | 2.17 | 0.63 | 0.05 |
| Number of students | 277 | | 310 | | 886 | | 887 | | | | | | |
| Percent missing data | ≤27% | | ≤22% | | ≤23% | | ≤20% | | | | | | |

[a]   Standard Deviations are only reported for Means or Mean Percents.

[b]   Estimated Impacts on Any Delinquency and No Delinquency students will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

[c]   Based on Benjamini-Hochberg test.

[d]   Lower Overall Absenteeism Rates indicate more positive outcomes.

[e]   Higher scores indicate higher grades; see Appendix F for further explanation of how these scores were derived.

*   p-value (of adjusted difference in means) < 0.05, two-tailed test.
+   p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing..

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit D.19: Subgroup Findings by Baseline Delinquency: Delinquent Behaviors and Participation in Harmful Activities**

| | Unadjusted Mean Outcome | | | | | | | | Estimated Impact | | | | |
| | Any Delinquency | | | | No Delinquency | | | | | | | | |
| | Treatment | | Control | | Treatment | | Control | | | | | | |
| Self-Reported Outcome[a] | Mean | Standard Deviation[b] | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation | Estimated Impact on Students Reporting Any Delinquency[c] | Estimated Impact on Students Reporting No Delinquency[c] | Difference in Impacts | p-value to Test Difference | BH-Corrected Critical Value[d] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Delinquency/Misconduct | 3.16 | 0.84 | 3.12 | 0.84 | 3.49 | 0.66 | 3.50 | 0.65 | 0.02 | 0.00 | 0.03 | 0.59 | 0.03 |
| Gang Activity | 12.26 | | 13.76 | | 3.81 | | 2.60 | | 0.57 | 1.18 | -0.61 | 0.83 | 0.04 |
| Tobacco Use | 15.91 | | 18.99 | | 3.12 | | 3.71 | | -3.11 | -0.54 | -2.56 | 0.44 | 0.03 |
| Alcohol Use | 29.14 | | 28.92 | | 6.62 | | 8.22 | | -0.93 | -1.72 | 0.78 | 0.84 | 0.05 |
| Drug Use | 17.97 | | 20.77 | | 4.06 | | 3.76 | | -2.83 | 0.09 | -2.92 | 0.37 | 0.02 |
| Number of students | 277 | | 310 | | 886 | | 887 | | | | | | |
| Percent missing data | ≤6% | | ≤6% | | ≤2% | | ≤3% | | | | | | |
| **School-Reported Behavioral Outcome** | | | | | | | | | | | | | |
| *Truancy Rate* [e,f] | 2.69 | 5.36 | 3.14 | 8.37 | 1.81 | 4.56 | 2.20 | 6.18 | -0.53 | -0.41 | -0.12 | 0.81 | 0.04 |
| Number of students | 277 | | 310 | | 886 | | 887 | | | | | | |
| Percent missing data | 38% | | 34% | | 44% | | 44% | | | | | | |
| *Misconduct* [f] | | | | | | | | | | | | | |
| Percent committing any infraction | 35.54 | | 28.42 | | 21.58 | | 21.06 | | 4.59 | 1.36 | 3.22 | 0.43 | 0.02 |
| Percent committing repeated infractions (2+) | 22.58 | | 18.48 | | 11.49 | | 14.68 | | 2.72 | -2.32 | 5.04 | 0.16 | 0.01 |
| *Delinquency* [f] | | | | | | | | | | | | | |
| Percent committing any infraction | 26.04 | | 32.13 | | 15.56 | | 15.97 | | -4.36 | -0.36 | -3.99 | 0.33 | 0.01 |
| Percent committing repeated infractions (2+) | 14.81 | | 16.12 | | 6.64 | | 6.79 | | -0.53 | -0.15 | -0.38 | 0.91 | 0.05 |
| Number of students | 277 | | 310 | | 886 | | 887 | | | | | | |
| Percent missing data | ≤18% | | ≤25% | | ≤22% | | ≤22% | | | | | | |

[a]   Higher scores on the Misconduct and Delinquency scales indicate more positive outcomes.

[b]   Standard Deviations are only reported for Means or Mean Percents.

[c]   Estimated Impacts on Any Delinquency and No Delinquency students will not necessarily be equal to the differences between the Unadjusted Mean Outcomes for these two groups.

[d]   Based on Benjamini-Hochberg test.

[e]   Based on 27 sites that reported unexcused absences and total days enrolled.

[f]   Lower percents of the school-reported Truancy, Misconduct, and  Delinquency items indicate more positive outcomes.

*   p-value (of adjusted difference in means) < 0.05, two-tailed test.
+   p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

# Appendix E: Sensitivity Tests

## Sensitivity of Impact Estimates to Nonresponse and Attrition

### Background

We obtained Student Survey data from students at all 42 Impact Study sites,[1] but student record data were unavailable for some sites. Furthermore, both Student Survey and school records data suffered from item nonresponse. If data were missing from our sample at random, there would be no reason to be concerned about nonresponse. On the other hand, if missing data in sites and/or for students occurred systematically, then our impact estimates would be subject to *selection bias* in how well they represent the full set of grantees and students chosen for the study.[2]

Suppose, for example, that—among those included in the study—the grantees with poor organization and leadership were less likely to be able to provide outcome data based on school records. If this poor organization and leadership were also associated with a poorly-run mentoring program with relatively small impacts, our estimated impacts would be biased upward, because the sample actually reporting outcome data would have contained disproportionate numbers of well-organized sites with strong leadership compared to the full set of 42 sites.

Similarly, at the student level, suppose that students in the treatment group who disliked school were less likely to answer questions on the Student Survey about Scholastic Efficacy and School Bonding. If these students (when in the treatment group) responded less positively to mentoring than students who enjoyed school, then our impact estimates would be biased upward, because our analytic sample would include a lower proportion of students who disliked school than the sample as a whole.

We therefore wished to examine the sensitivity of our impact estimates to nonresponse bias at both the site and the student levels. The remainder of this section describes our approach to comparing characteristics of students and sites with versus without missing data for each outcome of interest.

### Site-Level Nonresponse for School-reported Outcome Data

We were unable to collect school data on student absenteeism for 3 Impact Study sites; math, English language arts, and science grades for another 3 sites; social studies grades for 5 sites; state assessment test scores in math or reading/English language arts (ELA) for 1 site; truancy rate data for 15 sites; and disciplinary infractions data for 12 sites. We therefore wished to examine whether the sites with missing data appeared to differ systematically from sites that did report data for these outcomes, particularly for disciplinary infractions and truancy.

---

[1] Recall that, because some sites provided data in both years of our study, our sample consisted of 42 groups of students from 32 unique grantees.

[2] Note that, because our Impact Study sites were purposively selected, even if not subject to selection bias our impact estimates cannot be considered representative of impacts in the full population of 255 ED-funded Student Mentoring Programs. Rather, they characterize just the 32 programs from which the data are drawn. See Chapter 3 for a comparison of characteristics of Impact Study sites to characteristics for a representative sample of grantees.

Our first step was to examine, for each outcome measure of interest, baseline characteristics of students in sites not reporting data in comparison with baseline characteristics of students in sites with data for these measures. Because we had data for all sites from the baseline Student Survey, as well as school-reported demographic characteristics, including gender, age, minority status, family structure, and free- or reduced-price lunch eligibility for students in each site, we were able to calculate site-level means for each of these characteristics.

We then divided the sample into two subgroups for each outcome variable of interest: sites that reported data on the outcome category for some or all students, and sites that did not. Using the site-level weights used to perform the main impact analysis—i.e., weights set proportional to the number of students in each site—we then calculated weighted means for each subgroup. Finally, we performed a t test to determine whether differences in means across the two subgroups were statistically significant.[3,4]

The results of this exploratory analysis are reported below in Exhibit E.1. We found a number of statistically significant differences in observable student characteristics between sites with and without missing data:

- Sites not reporting truancy data had a lower proportion of boys than sites reporting truancy data.

- Sites missing data on absenteeism, state assessment tests, and math, English language arts, and science grades, had higher proportions of students aged 12 or older than sites reporting data for these outcomes, while sites not reporting data on truancy rates and social studies had lower proportions of older students.

- For all outcomes, sites with missing data had higher proportions of minority students than sites reporting outcome data for some or all students.

- Sites with missing data on disciplinary infractions had a lower proportion of students from two-parent families.

- Sites not reporting data on math, English language arts, and science grades, sites not reporting data on social studies grades, and sites not reporting data on academic proficiency from state assessment tests had higher proportions of students eligible for free- or reduced-price lunches.

---

[3]   Note that we did not control for multiple comparisons in this analysis. Of the sixty hypothesis tests performed here, one would expect to reject the null hypothesis in 3 cases due to random chance alone.

[4]   This exploratory analysis could test only for differences in observable student characteristics across sites. Even if no statistically significant differences in observable characteristics between sites with and without missing outcome data were detected, the possibility would remain that these groups of sites differed in characteristics not observable by the researcher that could also influence impacts—or that they differed to a modest degree on measured characteristics but the limited statistical power of the test procedure was unable to detect those differences.

**Exhibit E.1**

**Differences in Student Characteristics between Sites Reporting Outcome Data for Some or All Students and Sites with Missing Data for All Students**

| Site Missing Data on: | Overall Absenteeism Rate (3 missing sites) | | | Math, English Language Arts, and Science Grades (3 missing sites) | | | Social Studies Grades (5 missing sites) | | | State Assessment Tests, Math and Reading/E/LA (1 missing site) | | | Truancy Rate (15 missing sites) | | | Disciplinary Infractions (12 missing sites) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Sites With Outcome Data (n=2376) | Sites Without Outcome Data (n=197) | P-value to Test Difference | Sites With Outcome Data (n=2338) | Sites Without Outcome Data (n=235) | P-value to Test Difference | Sites With Outcome Data (n=2182) | Sites Without Outcome Data (n=391) | P-value to Test Difference | Sites With Outcome Data (n=2503) | Sites Without Outcome Data (n=70) | P-value to Test Difference | Sites With Outcome Data (n=1652) | Sites Without Outcome Data (n=921) | P-value to Test Difference | Sites With Outcome Data (n=1847) | Sites Without Outcome Data (n=726) | P-value to Test Difference |
| *Student Characteristics* | | | | | | | | | | | | | | | | | | |
| Proportion male | 0.48 | 0.42 | 0.08 | 0.47 | 0.48 | 0.81 | 0.47 | 0.51 | 0.07 | 0.47 | 0.50 | 0.65 | 0.51 | 0.41 | 0.00* | 0.48 | 0.45 | 0.21 |
| Proportion aged 12 and older | 0.29 | 0.36 | 0.01* | 0.29 | 0.38 | 0.00* | 0.30 | 0.26 | 0.03* | 0.29 | 0.46 | 0.00* | 0.32 | 0.27 | 0.00* | 0.28 | 0.34 | 0.00* |
| Proportion White | 0.23 | 0.01 | 0.00* | 0.23 | 0.06 | 0.00* | 0.22 | 0.17 | 0.00* | 0.22 | 0.02 | 0.00* | 0.26 | 0.14 | 0.00* | 0.23 | 0.18 | 0.00* |
| Proportion two-parent families | 0.56 | 0.59 | 0.46 | 0.56 | 0.58 | 0.59 | 0.56 | 0.59 | 0.26 | 0.56 | 0.59 | 0.69 | 0.56 | 0.58 | 0.32 | 0.55 | 0.61 | 0.00* |
| Proportion eligible for free or reduced-price lunch | 0.86 | 0.86 | 0.76 | 0.85 | 0.91 | 0.00* | 0.85 | 0.89 | 0.01* | 0.85 | 0.94 | 0.01* | 0.85 | 0.86 | 0.30 | 0.85 | 0.87 | 0.17 |
| *Baseline Student Survey Measures* | | | | | | | | | | | | | | | | | | |
| Pro-social Behaviors | 2.87 | 2.88 | 0.70 | 2.88 | 2.76 | 0.00* | 2.88 | 2.83 | 0.10 | 2.87 | 2.95 | 0.20 | 2.86 | 2.88 | 0.50 | 2.88 | 2.85 | 0.27 |
| Scholastic Efficacy and School Bonding | 3.15 | 3.25 | 0.01* | 3.16 | 3.12 | 0.27 | 3.16 | 3.13 | 0.42 | 3.15 | 3.21 | 0.35 | 3.12 | 3.22 | 0.00* | 3.17 | 3.10 | 0.00* |
| Future Orientation | 3.82 | 3.89 | 0.01* | 3.82 | 3.85 | 0.17 | 3.83 | 3.80 | 0.32 | 3.82 | 3.92 | 0.04* | 3.80 | 3.86 | 0.00* | 3.84 | 3.79 | 0.00* |
| Misconduct | 3.29 | 3.36 | 0.10 | 3.30 | 3.26 | 0.20 | 3.30 | 3.29 | 0.65 | 3.30 | 3.33 | 0.69 | 3.29 | 3.32 | 0.09 | 3.31 | 3.29 | 0.40 |
| Delinquency | 3.89 | 3.93 | 0.01* | 3.89 | 3.90 | 0.43 | 3.89 | 3.88 | 0.31 | 3.89 | 3.93 | 0.15 | 3.88 | 3.91 | 0.00* | 3.90 | 3.88 | 0.07 |

\* Difference was statistically significant at α=0.05, two-tailed test

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

- Students in sites not reporting grades in math, English language arts, and science exhibited lower scores on the Pro-social Behaviors scale from the Student Survey at baseline than students in sites reporting data for these outcomes.

- For absenteeism and truancy, students in sites with missing data responded more favorably at baseline on Student Survey scales measuring Scholastic Efficacy and School Bonding and Future Orientation, while students in sites with missing data on disciplinary infractions responded less favorably on these scales.

- Baseline student scores on the Delinquency scale from the Student Survey were higher in sites not reporting data on absenteeism rates and in sites not reporting data on truancy rates than in sites reporting these data.

The previous set of tests was able only to describe *baseline* differences between sites with and without missing data. While these results suggested that some differences between these subgroups of sites did exist, any conjecture about the degree of influence of these differences on *impacts* would necessarily be speculative.

However, because we had Student Survey outcome data for students in all sites, we were able to take our analysis one step further by examining whether impacts on survey-reported outcomes differed between sites with data on school-reported outcomes and all sites. We estimated impacts on our Student Survey outcomes both for the full Impact Study sample of 42 sites, and for the restricted subset of sites reporting data on each outcome measure for at least some or all students.[5] If the degree of impact estimate selection bias due to site nonresponse (which this procedure measures) is the same for Student Survey-reported outcomes as for school-reported outcomes, what we learn about the former will inform our understanding of the threat presented by the latter.

Each set of impact estimates was calculated as described in Chapter 2, but using a restricted set of sites rather than the full sample. For example, we estimated impacts on the Future Orientation scale from the Student Survey in sites providing data for some or all students on absentee rates.

We then conducted one-sample t tests to test whether the Future Orientation impact estimates in the restricted sample were different from Future Orientation impact estimates for the full set of sites, as presented in the main text. More specifically, the t values were calculated as:

$$[E.1] \quad t = \frac{\hat{\bar{\beta}}_1^* - \hat{\bar{\beta}}_1}{\hat{\sigma}_1^*},$$

---

[5] In order to estimate impacts, outcome data needed to be available for at least 2 treatment group and 2 control group students.

where

$\hat{\bar{\beta}}_1^*$ = estimated impact on subgroup of sites with school-reported outcome data on some or all students;

$\hat{\sigma}_1^*$ = estimated standard error of impact on subgroup of sites with school-reported outcome data for some or all students; and

$\hat{\bar{\beta}}_1$ = impact for all sites[6].

Exhibit E.2 presents impacts for the full sample of all sites compared to estimated impacts for the restricted sample of sites reporting data for each outcome of interest. There were no statistically significant differences in estimated impacts detected for any Student Survey outcomes. This finding increases our confidence that missing site-level data did not bias our estimated impacts on outcomes abstracted from student record data.

---

[6]  Although $\hat{\bar{\beta}}_1$ is an estimated impact, for the purpose of this analysis, we are treating it as a known "population" parameter (i.e., the true average impact for all 42 sites).   Hence, it does not have an associated error term.

**Exhibit E.2**

**Estimated Impact on Student Survey Outcomes in All Sites vs. Sites with Data on Outcome Variables from School Records**

| Student Survey Outcomes | Impact on All Sites | Sites with data on… | | | | | | | | | | | | |
| | | Overall Absenteeism Rate (3 missing sites) | | Math, English Language Arts, and Science Grades (3 missing sites) | | Social Studies Grades (5 missing sites) | | State Assessment Tests, Math and Reading/ELA (2 missing sites) | | Truancy Rate (15 missing sites) | | Disciplinary Infractions (12 missing sites) | |
| | | Impact on Sites with Data | P-value to Test Difference | Impact on Sites with Data | P-value to Test Difference | Impact on Sites with Data | P-value to Test Difference | Impact on Sites with Data | P-value to Test Difference | Impact on Sites with Data | P-value to Test Difference | Impact on Sites with Data | P-value to Test Difference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Pro-social Behaviors | -0.01 | 0.00 | 0.57 | -0.01 | 0.90 | -0.01 | 0.95 | 0.00 | 0.69 | 0.02 | 0.17 | 0.00 | 0.69 |
| Scholastic Efficacy & School Bonding | 0.04 | 0.03 | 0.78 | 0.03 | 0.71 | 0.02 | 0.32 | 0.04 | 0.96 | 0.02 | 0.59 | 0.04 | 0.80 |
| Future Orientation | 0.03 | 0.03 | 0.98 | 0.03 | 0.84 | 0.03 | 0.97 | 0.03 | 0.91 | 0.04 | 0.83 | 0.04 | 0.49 |
| Misconduct | 0.00 | 0.00 | 0.98 | 0.00 | 0.94 | 0.00 | 0.88 | 0.00 | 0.84 | 0.00 | 0.90 | -0.01 | 0.83 |
| Delinquency | 0.01 | 0.01 | 0.88 | 0.01 | 0.90 | 0.01 | 0.80 | 0.01 | 0.93 | 0.00 | 0.63 | 0.00 | 0.53 |

\* Difference was statistically significant at α=0.05, two-tailed test

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Student-Level Nonresponse**

In the previous subsection, we discussed potential bias in our impact estimates due to site-level nonresponse. However, even for those sites with school-reported outcome data, not all outcome measures were available for all students within each site. In this subsection, we discuss the sensitivity of our impact measures to this student-level nonresponse within the sites with school-reported outcome data.

Exhibit E.3 shows the range of item response rates by outcome measure for sites with school-reported outcome data on that measure for at least one student. Although, as reported in Chapter 2, response rates as a whole were quite high, it is evident from Exhibit E.3 that response rates for individual outcomes varied widely across sites, particularly for grades, state assessment test scores, and attendance measures. (Outcomes derived from student record data on disciplinary infractions were the exception; for these data, response rates were 100 percent for all sites reporting.)

**Exhibit E.3**

**Item Response Rates, Student Outcomes**

| Variable | Number Responding | | Response Rates in Sites Reporting Data | | |
|---|---|---|---|---|---|
| | Sites | Students | Total | Site-Level Minimum | Site-Level Maximum |
| *Student Survey Outcomes* | | | | | |
| Pro-social Behaviors | 42 | 2298 | 89.3% | 70.0% | 97.6% |
| Scholastic Efficacy & School Bonding | 42 | 2289 | 89.0% | 68.6% | 97.5% |
| Future Orientation | 42 | 2311 | 89.8% | 70.0% | 98.4% |
| Misconduct | 42 | 2329 | 90.5% | 70.0% | 100.0% |
| Delinquency | 42 | 2294 | 89.2% | 67.1% | 100.0% |
| *Grades* | | | | | |
| Math | 39 | 1677 | 71.7% | 9.4% | 100.0% |
| English Language Arts | 39 | 1692 | 72.4% | 8.2% | 100.0% |
| Science | 39 | 1633 | 69.8% | 11.8% | 100.0% |
| Social Studies | 37 | 1563 | 71.6% | 22.5% | 100.0% |
| *State Assessment Tests* | | | | | |
| Math | 41 | 1840 | 73.5% | 15.6% | 100.0% |
| Reading/ELA | 41 | 1837 | 73.4% | 25.2% | 100.0% |
| *Disciplinary Infractions* | | | | | |
| Misconduct | 30 | 1847 | 100.0% | 100.0% | 100.0% |
| Repeated Misconduct | 30 | 1847 | 100.0% | 100.0% | 100.0% |
| Delinquency | 30 | 1847 | 100.0% | 100.0% | 100.0% |
| Repeated Delinquency | 30 | 1847 | 100.0% | 100.0% | 100.0% |
| *Attendance* | | | | | |
| Overall Absenteeism Rate | 39 | 1978 | 83.2% | 17.5% | 100.0% |
| Truancy Rate | 27 | 1374 | 83.2% | 17.5% | 100.0% |

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

We therefore wished to examine whether, in sites providing outcome data for some, but not all, students, the students for whom outcome data were unavailable differed from students for whom these data were reported. We restricted the sample for this exploratory analysis to include only students from sites that had reported outcome data for at least one study participant. In this way, we hoped to isolate the phenomenon of *student-level* nonresponse from *site-level* nonresponse as examined in the previous subsection.

We then divided the sample into two subgroups for each outcome measure of interest: one subgroup comprised of students with available data for that outcome measure, and one comprised of students without available data.[7] Using site-level weights proportional to the number of students in each site, we then calculated weighted means of student baseline characteristics for each subgroup. Finally, we performed a t test to determine whether differences in means on student baseline characteristics across the two subgroups were statistically significant.[8]

This exploratory analysis could test only for differences in observable student characteristics across respondents and non-respondents. Note that our impact estimate regressions control for differences in these observable characteristics—our real concern is that *unobservable* differences between respondents and non-respondents may exist. Even if no statistically significant differences in observable characteristics between sites with and without missing outcome data were detected, the possibility would still remain that these groups of students differed in characteristics not observable by the researcher that could influence impact estimates—and we could not directly test that hypothesis. However, the presence of differences in observable characteristics between respondents and non-respondents could be an indicator that unobservable characteristics were also likely to differ between these groups.

The results of this analysis are presented in Exhibit E.4. We found statistically significant differences in response rates by age, race, and free or reduced-price lunch status. More specifically:

- There were no statistically significant differences in the proportion of boys among students with and without available data for either Student Survey outcomes or outcomes abstracted from school records.

- For 3 out of 5 Student Survey outcomes, students providing data were more likely to be aged 12 or older than students not providing data; students with available student record data were also more likely to be aged 12 or older for 5 out of 8 student record outcome measures.

- There were no differences in the proportion of White students with and without Student Survey data. However, students with data on grades from school records were more likely to be White than students without data on grades, and, conversely, students with data on truancy and absenteeism were less likely to be White than students without data on these attendance measures.

---

[7]  Note that we did not include outcome data from school records on disciplinary infractions in this analysis, because response rates were 100 percent for these items in all sites reporting any infractions.

[8]  Note that we did not control for multiple comparisons in this analysis. Of the 65 hypothesis tests performed here, one would expect to reject the null hypothesis in 3 to 4 cases due to random chance alone.

**Exhibit E.4**

**Item Nonresponse: Baseline Characteristics of Students in Sites Providing Outcome Data**

| Variable | Sites Reporting Data | Students in Sites Reporting Data | | Proportion Male | | | Proportion Aged 12 and Older | | | Proportion White | | | Proportion of Single-Parent Families | | | Proportion Eligible for Free or Reduced-Price Lunch | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Students with Outcome Data | Students without Outcome Data | Students with Outcome Data | Students without Outcome Data | P-value to Test Differ-ence | Students with Outcome Data | Students without Outcome Data | P-value to Test Differ-ence | Students with Outcome Data | Students without Outcome Data | P-value to Test Differ-ence | Students with Outcome Data | Students without Outcome Data | P-value to Test Differ-ence | Students with Outcome Data | Students without Outcome Data | P-value to Test Differ-ence |
| ***Student Survey Outcomes*** | | | | | | | | | | | | | | | | | | |
| Pro-social Behaviors | 42 | 2298 | 275 | 0.47 | 0.48 | 0.77 | 0.30 | 0.26 | 0.04* | 0.22 | 0.21 | 0.62 | 0.57 | 0.56 | 0.83 | 0.86 | 0.87 | 0.48 |
| Scholastic Efficacy & School Bonding | 42 | 2289 | 284 | 0.47 | 0.48 | 0.74 | 0.30 | 0.25 | 0.02* | 0.22 | 0.20 | 0.30 | 0.57 | 0.56 | 0.89 | 0.85 | 0.87 | 0.40 |
| Future Orientation | 42 | 2311 | 262 | 0.47 | 0.48 | 0.72 | 0.30 | 0.26 | 0.05 | 0.22 | 0.21 | 0.77 | 0.57 | 0.56 | 0.77 | 0.86 | 0.86 | 0.87 |
| Misconduct | 42 | 2329 | 244 | 0.47 | 0.48 | 0.73 | 0.30 | 0.26 | 0.07 | 0.22 | 0.21 | 0.79 | 0.57 | 0.56 | 0.93 | 0.86 | 0.86 | 0.58 |
| Delinquency | 42 | 2294 | 279 | 0.47 | 0.49 | 0.61 | 0.30 | 0.26 | 0.03* | 0.22 | 0.21 | 0.42 | 0.57 | 0.57 | 0.96 | 0.86 | 0.86 | 0.71 |
| ***Grades*** | | | | | | | | | | | | | | | | | | |
| Math | 39 | 1677 | 661 | 0.47 | 0.48 | 0.58 | 0.30 | 0.27 | 0.03* | 0.21 | 0.29 | 0.00* | 0.55 | 0.59 | 0.06 | 0.86 | 0.82 | 0.00* |
| English Language Arts | 39 | 1692 | 646 | 0.47 | 0.48 | 0.69 | 0.30 | 0.27 | 0.05 | 0.21 | 0.29 | 0.00* | 0.55 | 0.59 | 0.07 | 0.86 | 0.82 | 0.00* |
| Science | 39 | 1633 | 705 | 0.47 | 0.49 | 0.33 | 0.30 | 0.28 | 0.18 | 0.20 | 0.30 | 0.00* | 0.55 | 0.59 | 0.07 | 0.87 | 0.81 | 0.00* |
| Social Studies | 37 | 1563 | 619 | 0.46 | 0.48 | 0.50 | 0.31 | 0.30 | 0.82 | 0.20 | 0.28 | 0.00* | 0.55 | 0.59 | 0.07 | 0.87 | 0.81 | 0.00* |
| ***State Assessment Tests*** | | | | | | | | | | | | | | | | | | |
| Math | 41 | 1840 | 663 | 0.48 | 0.44 | 0.08 | 0.30 | 0.27 | 0.01* | 0.22 | 0.23 | 0.37 | 0.56 | 0.57 | 0.69 | 0.85 | 0.86 | 0.29 |
| Reading/ELA | 41 | 1837 | 666 | 0.48 | 0.44 | 0.07 | 0.30 | 0.26 | 0.01* | 0.22 | 0.23 | 0.57 | 0.56 | 0.57 | 0.66 | 0.85 | 0.86 | 0.53 |
| ***Attendance*** | | | | | | | | | | | | | | | | | | |
| Overall Absenteeism Rate | 39 | 1978 | 398 | 0.48 | 0.46 | 0.48 | 0.31 | 0.23 | 0.00* | 0.25 | 0.16 | 0.00* | 0.56 | 0.58 | 0.42 | 0.85 | 0.88 | 0.00* |
| Truancy Rate | 27 | 1374 | 278 | 0.51 | 0.51 | 0.99 | 0.33 | 0.26 | 0.00* | 0.28 | 0.17 | 0.00* | 0.55 | 0.59 | 0.20 | 0.84 | 0.90 | 0.00* |

\*    Difference was statistically significant at α=0.05, two-tailed test

*Source:*  Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

- There were no statistically significant differences in the proportion of students from two-parent families between respondents and non-respondents.

- Students without available data on grades from school records were less likely to be eligible for free or reduced-price lunches, while students without data on truancy and absenteeism were more likely to be eligible for free- or reduced-price lunches.

Because we found a number of statistically significant differences in student characteristics between respondents and non-respondents, we wished to determine the degree to which these differences were likely to have biased our impact estimates. Since our data came from a variety of sources, we were able to compare impact estimates on the full sample of students in sites reporting data with impacts based on the restricted sample of students reporting data for each outcome measure.

First, as above, for each outcome measure, we restricted the sample to only those sites reporting data on that measure.[9] For example, to determine the influence of item nonresponse for math grades, we restricted the sample to only the 39 sites reporting data on math grades for some or all students. This strategy allowed us to isolate the effects of student item nonresponse from the influence of site-level missing data as explained in the previous section. Using this restricted sample, we then estimated impacts on each of the other sixteen school-reported outcome and Student Survey-derived outcome measures for all students in these sites. To the extent that the degree of impact estimate selection bias due to student nonresponse is the same for survey-reported outcomes as for school-reported outcomes, what we learn about the former will inform our understanding of the threat presented by the latter.

Next, we further restricted the sample to include only students who reported data for the school-based outcome of interest, and again estimated the impacts on the other sixteen outcomes. For example, using the case of math grade data once again, we defined the sample to include only students from sites reporting math grades, for which we actually had math grades recorded in our dataset. We then estimated the impacts on Student Survey items, proficiency in math and reading/ELA based on state assessment tests, delinquency and misconduct outcomes from disciplinary infractions records, rates of truancy and absenteeism from attendance data, and grades in every subject except for math.

Finally, we compared the impact estimate based on only those students with data on the outcomes of interest to our originally reported impacts for all students (in sites reporting those outcomes) to determine the influence of item non-response for each school-reported outcome measure on impacts for all other outcome measures. Specifically, as described earlier, we performed a one-sample t test, where t was calculated as

$$[E.2] \quad t = \frac{\hat{\bar{\beta}}_1^a - \hat{\bar{\beta}}_1^b}{\hat{\sigma}_1^a},$$

where $\hat{\bar{\beta}}_1^a$ = estimated impact on students with data on the outcome of interest in sites reporting data on some or all students;

---

9    Note that we did not examine item nonresponse for outcomes derived from disciplinary infraction data in this analysis, because response rates were 100 percent for these items in all sites reporting any infractions.

$\hat{\sigma}_1^a$ = estimated standard error of impact on students with data on the outcome of interest in sites reporting data on some or all students; and

$\bar{\hat{\beta}}_1^b$ = impact on all students in sites reporting data on some or all students.[10]

Exhibits E.5-E.7 show the results of this exploratory analysis.[11] Student Survey item nonresponse was not associated with statistically significant changes in impact estimates for other outcome measures, as shown in Exhibit E.5. However, as seen in Exhibit E.6, excluding students for whom data on math, English language arts, and science grades were missing from the sample adversely affected impacts on delinquency and truancy rates from school records. Excluding students missing data on English language arts and science grades also improved impacts on disciplinary infractions due to misconduct. Finally, as shown in Exhibit E.7, item nonresponse for proficiency outcomes from state assessment tests and for attendance data on truancy and absenteeism, respectively, were not associated with any statistically significant differences in impacts between the restricted and full samples.

Overall, despite a few statistically significant differences in impacts, item nonresponse did not appear to greatly influence our impacts on other measures. When statistically differences were found, as in the case of missing student-level records on grades, the direction of the effect was inconsistent: some impacts were adversely affected when students missing data on grades were excluded, while others were improved. These results bolster our confidence that any systematic bias in our results arising from item nonresponse was likely small in magnitude and unlikely to influence the overall conclusions of our study.

**Student Survey Nonresponse**

In general, in evaluation studies researchers are also concerned with bias in impact estimates due to differential sample attrition in collecting outcome measures from follow-up surveys. If respondents lost to follow-up differ substantially from the full sample in terms of outcome levels or ability to benefit from the Student Mentoring Program, the resulting impact estimates will be biased.

However, in our study, Student Survey response rates were very high. As seen in Chapter 2, spring Student Survey response rates were above 92 percent for both treatment and control groups. Additionally, we did not find statistically significant differences in baseline student characteristics between students who did not complete a spring Student Survey and students who did. We therefore inferred that bias due to differential survey attrition was unlikely to be a major concern for the Impact Study.

---

[10]   Although $\bar{\hat{\beta}}_1^b$ is an estimated impact, for the purpose of this analysis, we are treating it as a known "population" parameter (i.e., the true average impact for all 42 sites). Hence, it does not have an associated error term.

[11]   We did not perform a correction for multiple comparisons in this analysis. Note that, of the 208 hypothesis tests we conducted, we would expect to find p-values below 0.05 for roughly 10 items due to random chance alone.

**Exhibit E.5: Item Nonresponse: Comparison of Impacts on All Students vs. Impacts on Students with Data on Student Survey Outcomes**

| Students Missing Student Survey Data on…. | Pro-social Behaviors | | | Scholastic Efficacy & School Bonding | | | Future Orientation | | | Misconduct | | | Delinquency | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Estimated Impacts on: | Impact on All Students | Impact on Students With Data | P-value to Test Difference | Impact on All Students | Impact on Students With Data | P-value to Test Difference | Impact on All Students | Impact on Students With Data | P-value to Test Difference | Impact on All Students | Impact on Students With Data | P-value to Test Difference | Impact on All Students | Impact on Students With Data | P-value to Test Difference |
| Pro-social Behaviors | | | | -0.01 | -0.01 | 0.91 | -0.01 | -0.01 | 0.87 | -0.01 | -0.01 | 0.92 | -0.01 | -0.01 | 0.89 |
| Scholastic Efficacy & School Bonding | 0.03 | 0.03 | 0.88 | | | | 0.03 | 0.03 | 0.92 | 0.03 | 0.03 | 0.88 | 0.03 | 0.03 | 0.96 |
| Future Orientation | 0.03 | 0.03 | 1.00 | 0.03 | 0.03 | 0.93 | | | | 0.03 | 0.03 | 0.98 | 0.03 | 0.03 | 0.89 |
| Misconduct | 0.00 | 0.00 | 0.93 | 0.00 | 0.01 | 0.88 | 0.00 | 0.00 | 0.88 | | | | 0.00 | 0.00 | 0.81 |
| Delinquency | 0.01 | 0.01 | 0.77 | 0.01 | 0.01 | 0.96 | 0.01 | 0.01 | 0.86 | 0.01 | 0.01 | 0.82 | | | |
| Math Grades | -0.08 | -0.10 | 0.85 | -0.08 | -0.11 | 0.81 | -0.08 | -0.11 | 0.87 | -0.08 | -0.11 | 0.85 | -0.08 | -0.10 | 0.98 |
| English Language Arts Grades | -0.05 | -0.01 | 1.00 | -0.05 | -0.02 | 0.81 | -0.05 | -0.01 | 0.88 | -0.05 | -0.01 | 0.99 | -0.05 | -0.01 | 0.96 |
| Science Grades | -0.03 | -0.08 | 0.66 | -0.03 | -0.09 | 0.58 | -0.03 | -0.09 | 0.61 | -0.03 | -0.09 | 0.52 | -0.03 | -0.08 | 0.70 |
| Social Studies Grades | -0.06 | -0.06 | 0.97 | -0.06 | -0.06 | 0.81 | -0.06 | -0.05 | 1.00 | -0.06 | -0.06 | 0.97 | -0.06 | -0.06 | 0.98 |
| Math Proficiency | 0.00 | 0.00 | 0.91 | 0.00 | -0.01 | 0.81 | 0.00 | 0.00 | 0.91 | 0.00 | -0.01 | 0.87 | 0.00 | 0.00 | 0.82 |
| Reading/ELA Proficiency | -0.03 | -0.02 | 0.87 | -0.03 | -0.02 | 0.95 | -0.03 | -0.02 | 0.85 | -0.03 | -0.02 | 0.76 | -0.03 | -0.03 | 0.98 |
| Misconduct | 0.00 | 0.00 | 0.87 | 0.00 | 0.00 | 0.96 | 0.00 | 0.00 | 0.74 | 0.00 | 0.00 | 0.81 | 0.00 | 0.00 | 0.99 |
| Repeated Misconduct | 0.00 | 0.00 | 0.32 | 0.00 | 0.00 | 0.58 | 0.00 | 0.00 | 0.43 | 0.00 | 0.00 | 0.42 | 0.00 | 0.00 | 0.61 |
| Delinquency | 0.03 | 0.03 | 0.87 | 0.03 | 0.04 | 0.83 | 0.03 | 0.04 | 0.76 | 0.03 | 0.04 | 0.79 | 0.03 | 0.04 | 0.60 |
| Repeated Delinquency | -0.01 | -0.01 | 0.93 | -0.01 | -0.01 | 0.97 | -0.01 | -0.01 | 0.95 | -0.01 | -0.01 | 0.94 | -0.01 | -0.01 | 0.99 |
| Overall Absenteeism Rate | -0.01 | -0.02 | 0.88 | -0.01 | -0.01 | 0.83 | -0.01 | -0.01 | 0.81 | -0.01 | -0.01 | 0.65 | -0.01 | -0.01 | 0.66 |
| Truancy Rate | -0.01 | -0.01 | 0.95 | -0.01 | 0.00 | 0.76 | -0.01 | 0.00 | 0.77 | -0.01 | 0.00 | 0.76 | -0.01 | 0.00 | 0.81 |

\* Difference was statistically significant at $\alpha=0.05$, two-tailed test

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit E.6: Item Nonresponse: Comparison of Impacts on All Students vs. Impacts on Students with Data on Grades from School Records**

| Students Missing Student Record Data on Grades in…. | Math | | | English Language Arts | | | Science | | | Social Studies | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Estimated Impacts on: | Impact on All Students | Impact on Students With Data | P-value to Test Difference | Impact on All Students | Impact on Students With Data | P-value to Test Difference | Impact on All Students | Impact on Students With Data | P-value to Test Difference | Impact on All Students | Impact on Students With Data | P-value to Test Difference |
| Pro-social Behaviors | -0.01 | -0.04 | 0.16 | -0.01 | -0.02 | 0.85 | -0.01 | -0.04 | 0.24 | -0.01 | -0.02 | 0.87 |
| Scholastic Efficacy & School Bonding | 0.02 | 0.01 | 0.73 | 0.02 | 0.00 | 0.55 | 0.02 | 0.02 | 0.63 | 0.01 | 0.02 | 0.43 |
| Future Orientation | 0.03 | 0.00 | 0.04* | 0.03 | 0.01 | 0.17 | 0.03 | 0.03 | 0.68 | 0.03 | 0.03 | 0.67 |
| Misconduct | 0.00 | -0.01 | 0.89 | 0.00 | -0.01 | 0.63 | 0.00 | -0.02 | 0.31 | 0.00 | -0.01 | 0.96 |
| Delinquency | 0.01 | 0.01 | 0.95 | 0.01 | 0.01 | 0.86 | 0.01 | 0.00 | 0.36 | 0.01 | 0.01 | 0.97 |
| Math Grades | | | | -0.08 | -0.07 | 0.89 | -0.08 | -0.08 | 0.88 | -0.06 | -0.05 | 0.92 |
| English Language Arts Grades | -0.05 | -0.08 | 0.76 | | | | -0.05 | -0.09 | 0.55 | -0.01 | -0.05 | 0.53 |
| Science Grades | -0.03 | -0.03 | 0.95 | -0.03 | -0.01 | 0.90 | | | | 0.00 | 0.00 | 0.90 |
| Social Studies Grades | -0.06 | -0.06 | 0.99 | -0.06 | -0.04 | 0.96 | -0.06 | -0.06 | 0.92 | | | |
| Math Proficiency | 0.00 | -0.01 | 0.97 | 0.00 | 0.00 | 0.90 | 0.00 | -0.01 | 0.95 | 0.00 | -0.02 | 0.46 |
| Reading/ELA Proficiency | -0.03 | -0.04 | 0.64 | -0.03 | -0.03 | 0.82 | -0.03 | -0.03 | 0.68 | -0.02 | -0.02 | 0.89 |
| Misconduct | 0.00 | 0.00 | 0.09 | 0.00 | -0.01 | 0.00* | 0.00 | 0.00 | 0.02* | 0.00 | 0.00 | 1.00 |
| Repeated Misconduct | 0.00 | 0.00 | 0.45 | 0.00 | 0.00 | 0.40 | 0.00 | 0.00 | 0.60 | 0.00 | 0.00 | 0.38 |
| Delinquency | 0.04 | 0.08 | 0.04* | 0.04 | 0.08 | 0.00* | 0.04 | 0.08 | 0.01* | 0.04 | 0.07 | 0.07 |
| Repeated Delinquency | -0.01 | 0.00 | 0.37 | -0.01 | 0.00 | 0.23 | -0.01 | 0.00 | 0.26 | -0.01 | -0.02 | 0.63 |
| Overall Absenteeism Rate | -0.02 | 0.01 | 0.10 | -0.02 | 0.00 | 0.07 | -0.02 | 0.01 | 0.08 | -0.02 | -0.01 | 0.62 |
| Truancy Rate | -0.01 | 0.03 | 0.00* | -0.01 | 0.02 | 0.00* | -0.01 | 0.03 | 0.00* | -0.01 | 0.01 | 0.06 |

\* Difference was statistically significant at α=0.05, two-tailed test

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit E.7: Item Nonresponse: Comparison of Impacts on All Students vs. Impacts on Students with Data on State Assessment Tests and Attendance from School Records**

| Students Missing Student Record Data on …. | School Records - State Assessment Test Scores | | | | | | School Records - Attendance | | | | | |
| | Math | | | Reading/ELA | | | Overall Absenteeism Rate | | | Truancy Rate | | |
| Estimated impacts on: | Impact on All Students | Impact on Students With Data | P-value to Test Difference | Impact on All Students | Impact on Students With Data | P-value to Test Difference | Impact on All Students | Impact on Students With Data | P-value to Test Difference | Impact on All Students | Impact on Students With Data | P-value to Test Difference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Pro-social Behaviors | -0.01 | -0.03 | 0.68 | -0.01 | -0.02 | 0.68 | 0.00 | -0.01 | 0.83 | 0.01 | 0.00 | 0.95 |
| Scholastic Efficacy & School Bonding | 0.03 | 0.03 | 0.77 | 0.03 | 0.03 | 0.80 | 0.03 | 0.04 | 0.51 | 0.01 | 0.02 | 0.56 |
| Future Orientation | 0.03 | 0.04 | 0.67 | 0.03 | 0.04 | 0.73 | 0.03 | 0.05 | 0.48 | 0.03 | 0.03 | 0.63 |
| Misconduct | 0.01 | 0.01 | 0.98 | 0.01 | 0.01 | 0.98 | 0.00 | 0.01 | 0.51 | 0.00 | 0.00 | 0.78 |
| Delinquency | 0.01 | 0.02 | 0.65 | 0.01 | 0.02 | 0.72 | 0.01 | 0.01 | 0.96 | 0.00 | 0.00 | 0.97 |
| Math Grades | -0.08 | -0.02 | 0.28 | -0.08 | -0.03 | 0.26 | -0.07 | -0.04 | 0.69 | -0.04 | 0.00 | 0.57 |
| English/Language Arts Grades | -0.05 | -0.03 | 0.99 | -0.05 | -0.04 | 0.89 | -0.06 | -0.05 | 0.81 | -0.04 | -0.04 | 0.84 |
| Science Grades | -0.03 | -0.04 | 0.73 | -0.03 | -0.04 | 0.83 | -0.03 | -0.02 | 0.87 | 0.02 | 0.02 | 0.97 |
| Social Studies Grades | -0.06 | -0.05 | 0.99 | -0.06 | -0.03 | 0.76 | -0.06 | -0.07 | 0.82 | -0.04 | -0.03 | 0.91 |
| Math Proficiency | | | | 0.00 | 0.00 | 0.94 | -0.01 | -0.02 | 0.51 | 0.00 | -0.01 | 0.52 |
| Reading/ELA Proficiency | -0.03 | -0.02 | 0.83 | | | | -0.03 | -0.04 | 0.58 | 0.00 | 0.00 | 0.79 |
| Misconduct | 0.00 | 0.00 | 0.58 | 0.00 | -0.01 | 0.28 | | | | 0.00 | 0.00 | 0.99 |
| Repeated Misconduct | 0.00 | 0.00 | 0.81 | 0.00 | 0.00 | 0.88 | 0.00 | 0.00 | 0.99 | | | |
| Delinquency | 0.03 | 0.02 | 0.79 | 0.03 | 0.02 | 0.83 | 0.03 | 0.05 | 0.36 | 0.03 | 0.05 | 0.24 |
| Repeated Delinquency | -0.01 | -0.02 | 0.76 | -0.01 | -0.02 | 0.76 | -0.01 | -0.01 | 0.72 | 0.00 | -0.01 | 0.70 |
| Overall Absenteeism Rate | -0.01 | -0.01 | 0.36 | -0.01 | -0.01 | 0.36 | -0.02 | -0.01 | 0.66 | -0.03 | -0.02 | 0.42 |
| Truancy Rate | -0.01 | -0.02 | 0.36 | -0.01 | -0.01 | 0.61 | -0.01 | -0.02 | 0.42 | -0.01 | -0.01 | 0.45 |

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

# Sensitivity of Impact Estimates to Alternative Weighting Methods

As explained in Chapter 2, we applied weights proportional to the sample size in each site when averaging site-level impact estimates to obtain the aggregate treatment effect. As a sensitivity test, we wished to examine the influence of our choice of weighting methodology on our estimates.

In this section of Appendix E, we therefore present a comparison of our main results to estimates based on three alternative weighting approaches: weighting by the total number of treatment students, weighting each site equally, and weighting by the inverse of the sampling variance. Each alternative methodology attaches a different relative importance to each site-level unit, reflecting a different conceptual framework for estimating the aggregate effect.

## Weights Proportional to the Site-Level Sample Size

We begin by reviewing the weighting methodology employed in calculating our main impact estimates, as first presented in Chapter 2: weighting each site proportionally to the total number of treatment and control group students in that site. Under this weighting scheme, the site-level weight $W_j^N$ for each site $j$ is defined as:

$$[E.5] \quad W_j = W_j^N = \frac{N_j}{\sum_{j=1}^{J} N_j}$$

Then the average impact estimate $\hat{\bar{\beta}}$ and its sampling variance $\hat{\sigma}^2$ are given by:

$$[E.6] \quad \hat{\bar{\beta}} = \sum_{j=1}^{J} W_j \hat{\beta}_j$$

$$[E.7] \quad \hat{\sigma}^2 = \sum_{j=1}^{J} W_j^2 \hat{\sigma}_j^2$$

where

$J$ = the total number of sites in the study;

$\hat{\beta}_j$ = the estimated treatment effect for site $j$;

$\hat{\sigma}_j^2$ = the estimated sampling variance for site $j$; and

$N_j$ = the total number of treatment and control students in site $j$.

Estimates incorporating this weighting methodology reflect the average treatment effect per student eligible to receive mentoring.

## Weights Proportional to the Site-Level Treatment Group Size

Suppose we were instead interested in determining the average treatment effect per student assigned to treatment. Intuitively, the appropriate weighting scheme would then weight sites proportionally to the size of the treatment group:

$$[E.8] \quad W_j = W_j^T = \frac{T_j}{\displaystyle\sum_{j=1}^{J} T_j},$$

where

$T_j$ = the number of students in the treatment group in site $j$.

The aggregate impact estimate and sampling variance are then calculated just as in equations [E.6] and [E.7] above, substituting the treatment group weight $W_j^T$ for the sample size weight $W_j^N$.

**No Site-Level Weights**

Next, suppose conceptually we wish to emphasize each site's contribution to the total impact estimate equally, regardless of its size. In this framework, we would weight each site proportionally—which is equivalent to taking the simple average across sites. The aggregate impact estimate and sampling variance are then given by:

$$[E.9] \quad \hat{\bar{\beta}} = \sum_{j=1}^{J} \hat{\beta}_j$$

$$[E.10] \quad \hat{\sigma}^2 = \sum_{j=1}^{J} \sigma_j^2$$

**Weights Proportional to the Inverse of the Sampling Variance**

Finally, we consider a fourth weighting method that is generally employed for statistical efficiency purposes. This method defines the weights proportionally to the inverse of the sampling variance for each site. By assigning greater importance to sites with less variation, this method provides an efficient estimator of the overall treatment effect (i.e., it has the lowest sampling variance of all possible impact estimators formed as weighted averages of the full set of site-specific impact estimates). That is:

$$[E.11] \quad W_j = W_j^V = \frac{1/\sigma_j^2}{\displaystyle\sum_{j=1}^{J} 1/\sigma_j^2}$$

As above, these site-level weights are substituted into equations [E.6] and [E.7] to obtain our estimate of the aggregate treatment effect.

**Comparison of Alternative Weighting Schemes**

Exhibits E.8 to E.10 present impact estimates calculated using each of these four weighting approaches. In general, weighting by the total sample size, weighting by the total treatment group size, and weighting by the inverse of the variance produce very similar results, since each of these three weighting methods attach greater importance to larger programs. The latter two weighting approaches yield findings of similar magnitude and the same level of statistical significance as the results weighted by the total sample size presented in the main analysis. In contrast, the method using uniform weights for sites attaches equal importance to each program, and yields no significant impacts. In other words, none of the null hypotheses are rejected in estimates incorporating uniform weights.

**Exhibit E.8: Estimated Impact on Interpersonal Relationships, Personal Responsibility, and Community Involvement by Different Weighting Methods**

| | Weights Set Proportional to | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | A. Total number of Treatment and Control Students in Site (Main Impact Findings) | | | B. Number of Students in Treatment Group | | | C. Uniform (Weights =1) | | | D. 1/Variance | | |
| Self-Reported Outcome | Impact | (SE) | P | Impact | (SE) | P | Impact | (SE) | P | Impact | (SE) | P |
| Pro-social Behaviors | -0.008 | (0.020) | 0.670 | -0.011 | (0.020) | 0.578 | -0.014 | (0.021) | 0.497 | -0.008 | (0.020) | 0.692 |

\* p-value (of adjusted difference in means) < 0.05, two-tailed test.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007.

**Exhibit E.9: Estimated Impact on Academic Outcomes by Different Weighting Methods**

| | Weights Set Proportional to | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | A. Total number of Treatment and Control Students in Site (Main Impact Findings) | | | B. Number of Students in Treatment Group | | | C. Uniform (Weights =1) | | | D. 1/Variance | | |
| | Impact | (SE) | P | Impact | (SE) | P | Impact | (SE) | P | Impact | (SE) | P |
| *Self-Reported Outcome (Scale Score: Range 1–4)* | | | | | | | | | | | | |
| Scholastic Efficacy and School Bonding | 0.036 | (0.021) | 0.077 | 0.039 | (0.021) | 0.059 | 0.029 | (0.022) | 0.178 | 0.037 | (0.021) | 0.074 |
| Future Orientation | 0.031* | (0.016) | 0.045 | 0.033* | (0.016) | 0.034 | 0.030 | (0.016) | 0.067 | 0.031* | (0.016) | 0.045 |
| *School-Reported Outcome* | | | | | | | | | | | | |
| *Overall Absenteeism Rate* | -0.463* | (0.228) | 0.042 | -0.445 | (0.231) | 0.054 | -0.281 | (0.267) | 0.294 | -0.466* | (0.228) | 0.041 |
| *Grades (Range 1–5)* | | | | | | | | | | | | |
| Math | -0.053 | (0.044) | 0.230 | -0.059 | (0.045) | 0.190 | -0.082 | (0.051) | 0.108 | -0.051 | (0.044) | 0.245 |
| English Language Arts | -0.037 | (0.044) | 0.397 | -0.049 | (0.045) | 0.271 | -0.048 | (0.053) | 0.360 | -0.038 | (0.044) | 0.393 |
| Science | -0.033 | (0.045) | 0.460 | -0.039 | (0.046) | 0.396 | -0.030 | (0.053) | 0.574 | -0.032 | (0.045) | 0.470 |
| Social Studies | -0.013 | (0.046) | 0.784 | -0.030 | (0.047) | 0.525 | -0.060 | (0.055) | 0.272 | -0.013 | (0.046) | 0.783 |
| *State Assessment Tests* | | | | | | | | | | | | |
| Math Proficiency | -1.534 | (1.851) | 0.407 | -1.936 | (1.889) | 0.305 | -0.495 | (2.190) | 0.821 | -1.667 | (1.849) | 0.367 |
| Reading/ELA Proficiency | -1.667 | (1.871) | 0.373 | -2.030 | (1.910) | 0.288 | -2.526 | (2.124) | 0.234 | -1.743 | (1.869) | 0.351 |

\* p-value (of adjusted difference in means) < 0.05, two-tailed test.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit E.10: Estimated Impact on Delinquent Behaviors and Participation in Harmful Activities by Different Weighting Methods**

| | Weights Set Proportional to | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | **A. Total Number of Treatment and Control Students in Site (Main Impact Findings)** | | | **B. Number of Students in Treatment Group** | | | **C. Uniform (Weights =1)** | | | **D. 1/Variance** | | |
| | **Impact** | **(SE)** | **P** | **Impact** | **(SE)** | **P** | **Impact** | **(SE)** | **P** | **Impact** | **(SE)** | **P** |
| *Self-Reported Behavioral Outcomes* | | | | | | | | | | | | |
| Misconduct | -0.001 | (0.021) | 0.954 | -0.003 | (0.021) | 0.874 | 0.005 | (0.022) | 0.835 | -0.001 | (0.021) | 0.962 |
| Delinquency | 0.010 | (0.009) | 0.285 | 0.009 | (0.009) | 0.321 | 0.011 | (0.010) | 0.280 | 0.010 | (0.009) | 0.284 |
| *School-Reported Behavioral Outcome* | | | | | | | | | | | | |
| *Truancy Rate* [a] | -0.450* | (0.200) | 0.024 | -0.428* | (0.202) | 0.034 | -0.425 | (0.225) | 0.059 | -0.450* | (0.200) | 0.024 |
| *Misconduct* | | | | | | | | | | | | |
| Percent committing any infraction | 2.560 | (1.688) | 0.129 | 2.330 | (1.703) | 0.171 | 3.311 | (1.767) | 0.061 | 2.542 | (1.688) | 0.132 |
| Percent committing repeated infractions (2+) | -0.981 | (1.400) | 0.483 | -0.857 | (1.412) | 0.544 | -1.077 | (1.465) | 0.462 | -1.004 | (1.400) | 0.473 |
| *Delinquency* | | | | | | | | | | | | |
| Percent committing any infraction | -1.510 | (1.627) | 0.353 | -1.637 | (1.641) | 0.319 | -1.239 | (1.703) | 0.467 | -1.504 | (1.627) | 0.355 |
| Percent committing repeated infractions (2+) | -0.559 | (1.215) | 0.645 | -0.532 | (1.225) | 0.664 | -0.765 | (1.271) | 0.547 | -0.571 | (1.215) | 0.638 |

[a]  Based on 27 sites that reported unexcused absences and total days enrolled

*  p-value (of adjusted difference in means) < 0.05, two-tailed test.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

## Sensitivity of Impact Estimates to Heteroscedasticity-Robust Standard Errors

In the context of regression, *heteroscedasticity* occurs when the variance of the error term is not constant—that is, in cases where different values of the dependent variable are associated with a wider or narrower distribution of error terms. The presence of heteroscedasticity violates one of the fundamental assumptions of linear regression modeling; failing to correct for it will lead to biased standard errors, invalidating the researcher's hypothesis tests. (Note that heteroscedasticity will *not* bias the point estimate of the treatment effect coefficient—only the standard errors of the estimate will be affected.)

In the context of the Student Mentoring Program, we were concerned about the possible presence of heteroscedasticity because of the wide range of observed baseline characteristics we saw across students. If, for example, students with greater tendencies toward delinquent behavior experience a wider variance in observed disciplinary infractions over time than students with less delinquency risk, our estimates could inappropriately reject the null hypothesis.

We therefore performed White's test to check for the presence of heteroscedasticity in our data. Recall from Chapter 2 that our WLS regression specification was given by:

$$[E.12] \quad \omega_{ij} Y_{ij} = \sum_{j=1}^{J} \omega_{ij}\, \beta_{1j} T_{ij}\, S_j + \sum_{j=1}^{J} \omega_{ij} \beta_{2j}\, S_j + \sum_{j=1}^{J} \omega_{ij} \beta_3\, X_{ij} + \omega_{ij}\, \varepsilon_{ij} ,$$

and our estimated variance by

$$[E.13] \quad Var(\hat{\beta}_{1j}) = \frac{\hat{\sigma}_j^{\,2}}{\sum_{i=1}^{N} (T_{ij} - \bar{T}_{ij})^2} \quad \text{for each given program } j = [1...J], \text{ where}$$

$$[E.14] \quad \hat{\sigma}_j^{\,2} = \sum_{i=1}^{N} \frac{\hat{\varepsilon}_{ij}^2}{N - K + 1}$$

$$[E.15] \quad \hat{\varepsilon}_{ij}^2 = (Y_i - \hat{\beta}_{1j} T_{ij} - \hat{\beta}_{2j} - \hat{\beta}_3 X_{ij})^2 ,$$

where,

$\omega_{ij}$ is the inverse of the probability of being randomly assigned to one's treatment status for student $i$ in program $j$,

$Y_{ij}$ is the outcome of interest $Y$ for student $i$ in program $j$,

$T_{ij}$ is the treatment indicator for student $i$ in program $j$ ($T_{ij} = 1$ if student $i$ is assigned to the treatment group; $T_{ij} = 0$ otherwise),

$S_j$ is a program indicator equal to 1 for students randomized at program $j$ and to 0 otherwise ($j = 1...J$),

$\beta_{1j}$ is the estimated average ITT treatment effect for program $j$,

$\beta_{2j}$ is the program fixed effect at program $j$ (i.e., the average untreated outcome level of a student at program $j$),

$X_{ij}$ is a vector of student characteristics measured for each student $i$ in program $j$,

$\beta_{3j}$ represents the vector of coefficients indicating how student characteristics affect student outcomes at program $j$, and

$\varepsilon_{ij}$ represents a random error term for student $i$ in program $j$, which is assumed to be independent and identically distributed across students.

We were concerned that the random error term was not in fact independent and identically distributed across students, in violation of one of the key assumptions of our model. To test for this possibility, we first estimated the above model in order to obtain the squared residual term, $\hat{\varepsilon}_{ij}^2$. We then regressed this term on the dependent variables from equation E.12 above, their squared values, and their cross products. White's statistic is then calculated as the product of the $R^2$ from this regression and the sample size $n$.

[E.16]   White's statistic $= n \cdot R^2$

This statistic has a chi-square distribution, with degrees of freedom equal to $k - 1$, where $k$ is the number of independent variables in the regression. Exhibit E.11 displays the values and associated p-values for White's statistic for each outcome in our study. As is evident from this table, we reject the null hypothesis that residuals were homoscedastic for 14 of our 17 outcome measures.

**Exhibit E.11**

**White's Test for Heteroscedasticity**

| Variable | White's Statistic | Chi-square Degrees of Freedom | P-value |
|---|---|---|---|
| *Interpersonal Relationships, Personal Responsibility, and Community Involvement* | | | |
| Pro-social Behaviors | 496.78 | 533 | 0.87 |
| ***Academic Outcome*** | | | |
| *Self-Reported Outcome* | | | |
| Scholastic Efficacy and School Bonding | 584.48 | 534 | 0.06 |
| Future Orientation | 749.71 | 533 | 0.00 |
| *School-Reported Outcome* | | | |
| Overall Absenteeism Rate | 1243.82 | 485 | 0.00 |
| Grades (Range 1–5) | | | |
| Math | 445.43 | 474 | 0.82 |
| English Language Arts | 494.41 | 467 | 0.18 |
| Science | 520.11 | 463 | 0.03 |
| Social Studies | 509.63 | 444 | 0.02 |
| State Assessment Tests | | | |
| Math Proficiency | 570.71 | 500 | 0.02 |
| Reading/ELA Proficiency | 637.98 | 503 | 0.00 |
| ***Delinquency Outcome*** | | | |
| *Self-Reported Outcome* | | | |
| Misconduct | 624.84 | 535 | 0.00 |
| Delinquency | 616.72 | 534 | 0.01 |
| *School-Reported Outcome* | | | |
| Truancy Rate | 693.36 | 345 | 0.00 |
| Misconduct | | | |
| Percent committing any infraction | 620.26 | 389 | 0.00 |
| Percent committing repeated infractions(2+) | 548.54 | 380 | 0.00 |
| Delinquency | | | |
| Percent committing any infraction | 597.20 | 392 | 0.00 |
| Percent committing repeated infractions(2+) | 615.32 | 370 | 0.00 |

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

We therefore employed the White method (White, 1980) to calculate heteroscedasticity-robust standard errors for our impact estimates. This method takes advantage of the fact that the squared residuals $\hat{\varepsilon}_{ij}^2$ represent a consistent estimate of the underlying unknown error variance $\sigma_{ij}^2$. The residuals are applied to adjust the variance term as follows:

$$[5] \quad Var_{robust}(\hat{\beta}_{1j}) = \frac{\sum_{i=1}^{N} r_{1ij}\hat{\varepsilon}_{ij}^2}{\left(\sum_{i=1}^{N}(T_{ij} - \overline{T}_j)^2\right)^2} \quad \text{for each given program } j = [1...J],$$

where $\hat{\varepsilon}_{ij}^2$ and $T_{ij}$ are defined as in the original model, and

$r_{1ij}$ = the $i$th residual from regressing $T_{ij}$ on all other dependent variables in the model.

Exhibit E.12 reports the resulting heteroscedasticity-robust standard errors and associated p-values alongside the non-robust standard errors and p-values for our impact estimates reported in the main text. In general, and consistently with our expectation, robust standard errors were slightly larger than the non-robust standard errors. However, the magnitude of the difference was not sufficient to influence overall statistical significance of our estimates.

**Exhibit E.12**

**Estimated Impacts, Comparison of Standard and Heteroscedasticity-Robust Standard Errors**

| | Unadjusted Mean Outcome | | Estimated Impact | | | | |
|---|---|---|---|---|---|---|---|
| | Treat-ment Group | Control Group | Impact on All Students | (SE) | P-value | (Robust SE) | Robust P-value |
| *Interpersonal Relationships, Personal Responsibility, and Community Involvement* | | | | | | | |
| Pro-social Behaviors | 2.79 | 2.80 | -0.01 | 0.02 | 0.67 | 0.02 | 0.68 |
| *Academic* | | | | | | | |
| *Self-Reported Outcome (Range 1–4):* | | | | | | | |
| Scholastic Efficacy and School Bonding | 3.06 | 3.03 | 0.04 | 0.02 | 0.08 | 0.02 | 0.08 |
| Future Orientation | 3.85 | 3.80 | 0.03*# | 0.02 | 0.05 | 0.02 | 0.05 |
| *School-Reported Outcome* | | | | | | | |
| Overall Absenteeism Rate | 5.04 | 5.52 | -0.46*# | 0.23 | 0.04 | 0.23 | 0.05 |
| Grades (Range 1–5) | | | | | | | |
| Math | 3.19 | 3.23 | -0.05 | 0.04 | 0.23 | 0.04 | 0.24 |
| English Language Arts | 3.57 | 3.61 | -0.04 | 0.04 | 0.40 | 0.05 | 0.41 |
| Science | 3.52 | 3.55 | -0.03 | 0.04 | 0.46 | 0.05 | 0.46 |
| Social Studies | 3.53 | 3.56 | -0.01 | 0.05 | 0.78 | 0.05 | 0.79 |
| State Assessment Tests | | | | | | | |
| Math Proficiency | 45.69 | 47.10 | -1.53 | 1.85 | 0.41 | 1.89 | 0.42 |
| Reading/ELA Proficiency | 49.40 | 50.76 | -1.67 | 1.87 | 0.37 | 1.89 | 0.38 |
| *Delinquent Behaviors and Participation in Harmful Activities* | | | | | | | |
| *Self-Reported Behavioral Outcome* | | | | | | | |
| Misconduct | 3.20 | 3.20 | 0.00 | 0.02 | 0.95 | 0.02 | 0.95 |
| Delinquency | 3.87 | 3.85 | 0.01 | 0.01 | 0.28 | 0.01 | 0.29 |
| School-Reported Behavioral Outcome | | | | | | | |
| Truancy Rate | 2.05 | 2.46 | -0.45*# | 0.20 | 0.02 | 0.20 | 0.03 |
| Misconduct | | | | | | | |
| Percent committing any infraction | 25.00 | 22.91 | 2.56 | 1.69 | 0.13 | 1.71 | 0.13 |
| Percent committing repeated infractions (2+) | 14.21 | 15.63 | -0.98 | 1.40 | 0.48 | 1.41 | 0.49 |
| Delinquency | | | | | | | |
| Percent committing any infraction | 18.13 | 20.03 | -1.51 | 1.63 | 0.35 | 1.65 | 0.36 |
| Percent committing repeated infractions (2+) | 8.64 | 9.13 | -0.55 | 1.21 | 0.65 | 1.23 | 0.65 |

\*    p-value (of adjusted difference in means) < 0.05, two-tailed test.

\#    Robust p-value (of adjusted difference in means) < 0.05, two-tailed test.

*Source:  Source:*  Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

# Appendix F: Standard Errors and Confidence Intervals of Main Effects

**Exhibit F.1**

**Standard Errors and Confidence Intervals for Estimated Impact on Interpersonal Relationships, Personal Responsibility, and Community Involvement**

| Self-Reported Outcome | Estimated Impact | Standard Error | 95% Confidence Interval |
|---|---|---|---|
| Pro-social Behaviors | -0.01 | (0.02) | -0.05 — 0.03 |

\*   p-value (of adjusted difference in means) < 0.05, two-tailed test.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006and Spring 2007.

**Exhibit F.2**

**Standard Errors and Confidence Intervals for Estimated Impacts on Academic Outcomes**

| Self-Reported Outcome | Estimated Impact | Standard Error | 95% Confidence Interval |
|---|---|---|---|
| Scholastic Efficacy and School Bonding | 0.04 | (0.02) | 0.00 — 0.08 |
| Future Orientation | 0.03* | (0.02) | 0.00 — 0.06 |
| **School-Reported Outcome** | | | |
| *Overall Absenteeism Rate (Percent)* | -0.46* | (0.23) | -0.91 — -0.02 |
| *Grades (Range 1–5)* | | | |
| Math | -0.05 | (0.04) | -0.14 — 0.03 |
| English Language Arts | -0.04 | (0.04) | -0.12 — 0.05 |
| Science | -0.03 | (0.04) | -0.12 — 0.06 |
| Social Studies | -0.01 | (0.05) | -0.10 — 0.08 |
| *State Assessment Tests* | | | |
| Math—Percent Proficient | -1.53 | (1.85) | -5.16 — 2.10 |
| Reading/ELA—Percent Proficient | -1.67 | (1.87) | -5.33 — 2.00 |

\*   p-value (of adjusted difference in means) < 0.05, two-tailed test.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit F.3**

**Standard Errors and Confidence Intervals for Estimated Impacts on Delinquent Behaviors and Participation in Harmful Activities**

| Self-Reported Outcome | Estimated Impact | Standard Error | 95% Confidence Interval |
|---|---|---|---|
| Misconduct | 0.00 | (0.02) | -0.04 — 0.04 |
| Delinquency | 0.01 | (0.01) | -0.01 — 0.03 |
| **School-Reported Behavioral Outcome** | | | |
| *Truancy—Unexcused Absence Rate (Percent)* [a] | -0.45* | (0.20) | -0.84 — -0.06 |
| *Misconduct* | | | |
| Percent committing any infraction | 2.56 | (1.69) | -0.75 — 5.87 |
| Percent committing repeated infractions (2+) | -0.98 | (1.40) | -3.72 — 1.76 |
| *Delinquency* | | | |
| Percent committing any infraction | -1.51 | (1.63) | -4.70 — 1.68 |
| Percent committing repeated infractions (2+) | -0.56 | (1.21) | -2.94 — 1.82 |

[a]  Based on 27 of 42 sites that reported unexcused absences and total days enrolled

\*   p-value (of adjusted difference in means) < 0.05, two-tailed test

*Source:*  Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit F.4: Subgroup Findings by Gender: Interpersonal Relationships, Personal Responsibility, and Community Involvement**

| Self-Reported Outcome | Boys | | | Girls | | | Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| | Estimated Impact | Standard Error | 95% Confidence Interval | Estimated Impact | Standard Error | 95% Confidence Interval | Estimated Impact | Standard Error | 95% Confidence Interval |
| Pro-social Behaviors | -0.06*+ | (0.03) | -0.12 — 0.00 | 0.04 | (0.03) | -0.01 — 0.09 | -0.10* | (0.04) | -0.18 — -0.03 |

\* p-value (of adjusted difference in means) < 0.05, two-tailed test.

+ p-value (of adjusted difference in means) < BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006and Spring 2007.

### Exhibit F.5: Subgroup Findings by Gender: Academic Outcomes

| Self-Reported Outcome (Range 1–4) | Boys | | | Girls | | | Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| | Estimated Impact | Standard Error | 95% Confidence Interval | Estimated Impact | Standard Error | 95% Confidence Interval | Estimated Impact | Standard Error | 95% Confidence Interval |
| Scholastic Efficacy and School Bonding | -0.03 | (0.03) | -0.09 — 0.03 | 0.10*+ | (0.03) | 0.04 — 0.15 | -0.12*+ | (0.04) | -0.21 — -0.04 |
| Future Orientation | 0.07*+ | (0.02) | 0.02 — 0.12 | 0.00 | (0.02) | -0.04 — 0.04 | 0.07*+ | (0.03) | 0.01 — 0.13 |
| **School-Reported Outcome** | | | | | | | | | |
| ***Overall Absenteeism Rate (Percent)*** | -0.12 | (0.30) | -0.72 — 0.47 | -0.82* | (0.34) | -1.49 — -0.14 | 0.69 | (0.46) | -0.21 — 1.59 |
| *Grades (Range 1–5)* | | | | | | | | | |
| Math | -0.07 | (0.07) | -0.20 — 0.06 | -0.04 | (0.06) | -0.16 — 0.08 | -0.03 | (0.09) | -0.21 — 0.15 |
| English Language Arts | -0.08 | (0.07) | -0.21 — 0.05 | 0.05 | (0.06) | -0.07 — 0.17 | -0.12 | (0.09) | -0.31 — 0.05 |
| Science | -0.04 | (0.07) | -0.17 — 0.09 | -0.03 | (0.06) | -0.15 — 0.10 | -0.01 | (0.09) | -0.20 — 0.17 |
| Social Studies | 0.01 | (0.07) | -0.13 — 0.15 | -0.05 | (0.06) | -0.18 — 0.07 | 0.06 | (0.09) | -0.13 — 0.24 |
| ***State Assessment Tests*** | | | | | | | | | |
| Math—Percent Proficient | -2.38 | (2.74) | -7.75 — 2.98 | -1.17 | (2.61) | -6.28 — 3.94 | -1.28 | (3.78) | -8.62 — 6.20 |
| Reading/ELA—Percent Proficient | 1.61 | (2.69) | -3.67 — 6.88 | -4.91 | (2.66) | -10.12 — 0.30 | 6.52 | (3.78) | -0.89 — 13.93 |

\* p-value (of adjusted difference in means) < 0.05, two-tailed test.
\+ p-value (of adjusted difference in means) < BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit F.6: Subgroup Findings by Gender: Delinquent Behaviors and Participation in Harmful Activities**

| Self-Reported Outcome | Boys | | | Girls | | | Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| | Estimated Impact | Standard Error | 95% Confidence Interval | Estimated Impact | Standard Error | 95% Confidence Interval | Estimated Impact | Standard Error | 95% Confidence Interval |
| Misconduct | 0.02 | (0.03) | -0.04 — 0.09 | -0.01 | (0.03) | -0.07 — 0.04 | 0.04 | (0.04) | -0.05 — 0.12 |
| Delinquency | 0.01 | (0.02) | -0.02 — 0.04 | 0.01 | (0.01) | -0.01 — 0.03 | -0.01 | (0.02) | -0.04 — 0.03 |
| **School-Reported Behavioral Outcome** | | | | | | | | | |
| ***Truancy—Unexcused Absence Rate (Percent)[a]*** | -0.23 | (0.24) | -0.70 — 0.24 | -0.76* | (0.33) | -1.40 — -0.11 | 0.53 | (0.41) | -0.27 — 1.33 |
| ***Misconduct*** | | | | | | | | | |
| Percent committing any infraction | 6.33 | (2.65) | 1.14 — 11.52 | -1.21 | (2.12) | -5.37 — 2.95 | 7.54* | (3.39) | 0.89 — 14.19 |
| Percent committing repeated infractions (2+) | 0.83 | (2.19) | -3.46 — 5.11 | -2.55 | (1.79) | -6.07 — 0.97 | 3.38 | (2.83) | -2.17 — 8.93 |
| ***Delinquency*** | | | | | | | | | |
| Percent committing any infraction | 0.16 | (2.66) | -7.01 — 7.34 | -2.55 | (2.04) | -6.55 — 1.45 | 2.71 | (3.30) | -3.77 — 9.19 |
| Percent committing repeated infractions (2+) | 0.61 | (3.04) | -5.35 — 6.57 | -0.97 | (1.52) | -3.95 — 2.00 | 1.59 | (2.46) | -3.24 — 6.41 |

[a]   Based on 27 sites that reported unexcused absences and total days enrolled.

\*   p-value (of adjusted difference in means) < 0.05, two-tailed test.
+   p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing

*Source:*  Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit F.7: Subgroup Findings by Age: Interpersonal Relationships, Personal Responsibility, and Community Involvement**

| Self-Reported Outcome | Students Below Age 12 | | | Students Aged 12 and Older | | | Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| | Estimated Impact | Standard Error | 95% Confidence Interval | Estimated Impact | Standard Error | 95% Confidence Interval | Estimated Impact | Standard Error | 95% Confidence Interval |
| Pro-social Behaviors | -0.02 | (0.02) | -0.07 — 0.03 | 0.03 | (0.03) | -0.04 — 0.10 | -0.05 | (0.04) | -0.13 — 0.03 |

* p-value (of adjusted difference in means) < 0.05, two-tailed test.

+ p-value (of adjusted difference in means) < BH-Corrected Critical Value ➜ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006and Spring 2007.

**Exhibit F.8: Subgroup Findings by Age: Academic Outcomes**

| Self-Reported Outcome | Students Below Age 12 | | | Students Aged 12 and Older | | | Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| | Estimated Impact | Standard Error | 95% Confidence Interval | Estimated Impact | Standard Error | 95% Confidence Interval | Estimated Impact | Standard Error | 95% Confidence Interval |
| Scholastic Efficacy and School Bonding | 0.02 | (0.02) | -0.03 — 0.07 | 0.08* | (0.04) | 0.00 — 0.16 | -0.06 | (0.05) | -0.15 — 0.03 |
| Future Orientation | 0.03 | (0.02) | 0.00 — 0.07 | 0.04 | (0.03) | -0.02 — 0.10 | -0.01 | (0.04) | -0.08 — 0.06 |
| **School-Reported Outcome** | | | | | | | | | |
| ***Overall Absenteeism Rate (Percent)*** | -0.51* | (0.21) | -0.93 — -0.10 | -0.88 | (0.56) | -1.98 — 0.23 | 0.37 | (0.60) | -0.81 — 1.55 |
| ***Grades (Range 1–5)*** | | | | | | | | | |
| Math | -0.07 | (0.05) | -0.17 — 0.04 | -0.03 | (0.10) | -0.22 — 0.17 | -0.04 | (0.11) | -0.26 — 0.18 |
| English Language Arts | -0.04 | (0.05) | -0.14 — 0.07 | -0.05 | (0.09) | -0.22 — 0.12 | 0.01 | (0.10) | -0.19 — 0.22 |
| Science | -0.02 | (0.06) | -0.13 — 0.09 | -0.02 | (0.08) | -0.18 — 0.14 | 0.00 | (0.10) | -0.20 — 0.19 |
| Social Studies | 0.02 | (0.06) | -0.10 — 0.13 | -0.01 | (0.09) | -0.18 — 0.15 | 0.03 | (0.10) | -0.17 — 0.23 |
| ***State Assessment Tests*** | | | | | | | | | |
| Math—Percent Proficient | -3.24 | (2.25) | -7.65 — 1.17 | 1.13 | (3.33) | -5.40 — 7.66 | -4.37 | (4.02) | -12.25 — 3.51 |
| Reading/ELA—Percent Proficient | -1.83 | (2.23) | -6.20 — 2.54 | -1.12 | (3.50) | -7.98 — 5.74 | -0.71 | (4.15) | -8.85 — 7.43 |

\*   p-value (of adjusted difference in means) < 0.05, two-tailed test.

+   p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing.

*Source:*  Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit F.9: Subgroup Findings by Age: Delinquent Behaviors and Participation in Harmful Activities**

| Self-Reported Outcome | Students Below Age 12 | | | Students Aged 12 and Older | | | Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| | Estimated Impact | Standard Error | 95% Confidence Interval | Estimated Impact | Standard Error | 95% Confidence Interval | Estimated Impact | Standard Error | 95% Confidence Interval |
| Misconduct | 0.01 | (0.02) | -0.03 — 0.06 | -0.04 | (0.04) | -0.12 — 0.04 | 0.06 | (0.05) | -0.04 — 0.15 |
| Delinquency | 0.01 | (0.01) | -0.01 — 0.03 | 0.02 | (0.02) | -0.02 — 0.06 | -0.01 | (0.02) | -0.06 — 0.03 |
| **School-Reported Behavioral Outcome** | | | | | | | | | |
| ***Truancy—Unexcused Absence Rate (Percent)[a]*** | -0.55*+ | (0.18) | -0.91 — -0.19 | -0.84 | (0.48) | -1.78 — 0.09 | 0.30 | (0.51) | -0.71 — 1.30 |
| ***Misconduct*** | | | | | | | | | |
| Percent committing any infraction | 4.11* | (1.96) | 0.26 — 7.96 | -3.03 | (3.20) | -9.30 — 3.24 | 7.14 | (3.75) | -0.21 — 14.50 |
| Percent committing repeated infractions (2+) | 0.46 | (1.57) | -2.61 — 3.52 | -6.32* | (2.79) | -11.79 — -0.86 | 6.78* | (3.20) | 0.51 — 13.04 |
| ***Delinquency*** | | | | | | | | | |
| Percent committing any infraction | -0.22 | (1.83) | -3.82 — 3.37 | -6.40 | (3.48) | -13.21 — 0.42 | 6.17 | (3.93) | -1.53 — 13.88 |
| Percent committing repeated infractions (2+) | -0.56 | (1.19) | -2.90 — 1.77 | -1.09 | (3.05) | -7.07 — 4.88 | 0.53 | (3.27) | -5.88 — 6.95 |

[a]  Based on 27 sites that reported unexcused absences and total days enrolled.

\*  p-value (of adjusted difference in means) < 0.05, two-tailed test.
+  p-value (of adjusted difference in means) < BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit F.10: Subgroup Findings by Family Structure: Interpersonal Relationships, Personal Responsibility, and Community Involvement**

| Self-Reported Outcome | Two-Parent Households | | | Other Households | | | Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| | Estimated Impact | Standard Error | 95% Confidence Interval | Estimated Impact | Standard Error | 95% Confidence Interval | Estimated Impact | Standard Error | 95% Confidence Interval |
| Pro-social Behaviors | 0.02 | (0.03) | -0.04 — 0.07 | -0.04 | (0.03) | -0.10 — 0.02 | -0.06 | (0.04) | -0.02 — 0.14 |

[*]  p-value (of adjusted difference in means) < 0.05, two-tailed test.

[+]  p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006and Spring 2007.

**Exhibit F.11: Subgroup Findings by Family Structure: Academic Outcomes**

| | Two-Parent Households | | | Other Households | | | Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| **Self-Reported Outcome** | **Estimated Impact** | **Standard Error** | **95% Confidence Interval** | **Estimated Impact** | **Standard Error** | **95% Confidence Interval** | **Estimated Impact** | **Standard Error** | **95% Confidence Interval** |
| Scholastic Efficacy and School Bonding | 0.03 | (0.03) | -0.03 — 0.08 | 0.05 | (0.03) | -0.02 — 0.11 | -0.02 | (0.04) | -0.10 — 0.06 |
| Future Orientation | 0.03 | (0.02) | -0.01 — 0.07 | 0.03 | (0.02) | -0.02 — 0.08 | 0.01 | (0.03) | -0.06 — 0.07 |
| **School-Reported Outcome** | | | | | | | | | |
| ***Overall Absenteeism Rate (Percent)*** | -0.19 | (0.25) | -0.68 — 0.30 | -0.65 | (0.40) | -1.44 — 0.13 | 0.46 | (0.47) | -0.46 — 1.39 |
| *Grades (Range 1–5)* | | | | | | | | | |
| Math | -0.09 | (0.06) | -0.21 — 0.02 | -0.01 | (0.07) | -0.14 — 0.13 | -0.09 | (0.09) | -0.27 — 0.09 |
| English Language Arts | -0.03 | (0.06) | -0.15 — 0.09 | -0.02 | (0.07) | -0.15 — 0.12 | -0.01 | (0.09) | -0.19 — 0.16 |
| Science | -0.04 | (0.06) | -0.16 — 0.08 | -0.05 | (0.07) | -0.18 — 0.09 | 0.01 | (0.09) | -0.17 — 0.19 |
| Social Studies | 0.04 | (0.06) | -0.09 — 0.16 | 0.02 | (0.07) | -0.12 — 0.16 | 0.02 | (0.10) | -0.17 — 0.21 |
| ***State Assessment Tests*** | | | | | | | | | |
| Math—Percent Proficient | -2.38 | (2.74) | -7.75 — 2.98 | -1.17 | (2.61) | -6.28 — 3.94 | -1.21 | (3.78) | -8.62 — 6.20 |
| Reading/ELA—Percent Proficient | 1.61 | (2.69) | -3.67 — 6.88 | -4.91 | (2.66) | -10.12 — 0.30 | 6.52 | (3.78) | -0.89 — 13.93 |

\*    p-value (of adjusted difference in means) < 0.05, two-tailed test.
+    p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing.

*Source:*  Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit F.12: Subgroup Findings by Family Structure: Delinquent Behaviors and Participation in Harmful Activities**

| Self-Reported Outcome | Two-Parent Households | | | Other Households | | | Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| | Estimated Impact | Standard Error | 95% Confidence Interval | Estimated Impact | Standard Error | 95% Confidence Interval | Estimated Impact | Standard Error | 95% Confidence Interval |
| Misconduct | -0.01 | (0.03) | -0.06 — 0.05 | 0.00 | (0.03) | -0.06 — 0.06 | -0.01 | (0.04) | -0.09 — 0.08 |
| Delinquency | 0.01 | (0.01) | -0.01 — 0.04 | 0.00 | (0.01) | -0.02 — 0.03 | 0.01 | (0.02) | -0.03 — 0.05 |
| **School-Reported Behavioral Outcome** | | | | | | | | | |
| *Truancy—Unexcused Absence Rate (Percent)*[a] | 0.02 | (0.20) | -0.36 — 0.41 | -0.90 | (0.37) | -1.63 — -0.16 | 0.92 | (0.42) | 0.09 — 1.75 |
| *Misconduct* | | | | | | | | | |
| Percent committing any infraction | 1.76 | (2.23) | -2.62 — 6.14 | 2.83 | (2.67) | -2.39 — 8.06 | -1.07 | (3.48) | -7.89 — 5.74 |
| Percent committing repeated infractions (2+) | -2.16 | (1.80) | -5.68 — 1.37 | 0.30 | (2.24) | -4.09 — 4.69 | -2.46 | (2.87) | -8.08 — 3.17 |
| *Delinquency* | | | | | | | | | |
| Percent committing any infraction | -3.56 | (3.66) | -10.73 — 3.62 | -0.06 | (2.64) | -5.24 — 5.12 | -3.50 | (3.36) | -10.08 — 3.09 |
| Percent committing repeated infractions (2+) | -1.92 | (3.04) | -7.89 — 4.04 | 0.15 | (2.04) | -3.85 — 4.15 | -2.08 | (2.52) | -7.02 — 2.86 |

[a]    Based on 27 sites that reported unexcused absences and total days enrolled.

\*    p-value (of adjusted difference in means) < 0.05, two-tailed test.
+    p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing

*Source:*  Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit F.13: Subgroup Findings by Academic Risk: Interpersonal Relationships, Personal Responsibility, and Community Involvement**

| | Proficient Students | | | Not Proficient Students | | | Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| **Self-Reported Outcome** | **Estimated Impact** | **Standard Error** | **95% Confidence Interval** | **Estimated Impact** | **Standard Error** | **95% Confidence Interval** | **Estimated Impact** | **Standard Error** | **95% Confidence Interval** |
| Pro-social Behaviors | -0.01 | (0.03) | -0.08 — 0.06 | -0.01 | (0.03) | -0.07 — 0.05 | 0.00 | (0.04) | -0.09 — 0.09 |

\* p-value (of adjusted difference in means) < 0.05, two-tailed test.

+ p-value (of adjusted difference in means) < BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006and Spring 2007.

**Exhibit F.14: Subgroup Findings by Academic Risk: Academic Outcomes**

| | Proficient Students | | | Not Proficient Students | | | Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| **Self-Reported Outcome** | **Estimated Impact** | **Standard Error** | **95% Confidence Interval** | **Estimated Impact** | **Standard Error** | **95% Confidence Interval** | **Estimated Impact** | **Standard Error** | **95% Confidence Interval** |
| Scholastic Efficacy and School Bonding | 0.01 | (0.03) | -0.06 — 0.08 | 0.05 | (0.03) | -0.01 — 0.11 | -0.04 | (0.05) | -0.13 — 0.05 |
| Future Orientation | 0.03 | (0.02) | -0.01 — 0.08 | 0.04 | (0.02) | 0.00 — 0.09 | -0.01 | (0.03) | -0.08 — 0.05 |
| **School-Reported Outcome** | | | | | | | | | |
| *Overall Absenteeism Rate (Percent)* | -0.60 | (0.36) | -1.29 — -0.10 | -0.72* | (0.32) | -1.34 — -0.10 | 0.13 | (0.48) | -0.81 — 1.06 |
| *Grades (Range 1–5)* | | | | | | | | | |
| Math | -0.03 | (0.08) | -0.18 — 0.13 | -0.02 | (0.07) | -0.15 — 0.11 | 0.00 | (0.10) | -0.20 — 0.20 |
| English Language Arts | 0.01 | (0.08) | -0.14 — 0.16 | -0.07 | (0.07) | -0.20 — 0.06 | 0.09 | (0.10) | -0.11 — 0.29 |
| Science | -0.08 | (0.08) | -0.24 — 0.07 | -0.04 | (0.07) | -0.17 — 0.09 | -0.05 | (0.10) | -0.25 — 0.16 |
| Social Studies | 0.07 | (0.08) | -0.09 — 0.23 | 0.00 | (0.07) | -0.14 — 0.13 | 0.07 | (0.11) | -0.14 — 0.28 |
| *State Assessment Tests* | | | | | | | | | |
| Math—Percent Proficient | -3.35 | (2.94) | -9.11 — 2.42 | 0.33 | (2.39) | -4.36 — 5.02 | -3.68 | (3.79) | -11.11 — 3.75 |
| Reading/ELA—Percent Proficient | -1.75 | (2.73) | -7.11 — 3.60 | -0.53 | (2.53) | -5.49 — 4.44 | -1.23 | (3.73) | -8.53 — 6.08 |

\* p-value (of adjusted difference in means) < 0.05, two-tailed test.

+ p-value (of adjusted difference in means) < BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit F.15: Subgroup Findings by Academic Risk: Delinquent Behaviors and Participation in Harmful Activities**

| Self-Reported Outcome | Proficient Students | | | Not Proficient Students | | | Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| | Estimated Impact | Standard Error | 95% Confidence Interval | Estimated Impact | Standard Error | 95% Confidence Interval | Estimated Impact | Standard Error | 95% Confidence Interval |
| Misconduct | 0.02 | (0.03) | -0.05 — 0.09 | -0.02 | (0.03) | -0.08 — 0.04 | 0.04 | (0.05) | -0.05 — 0.13 |
| Delinquency | 0.01 | (0.02) | -0.03 — 0.04 | 0.01 | (0.01) | -0.02 — 0.04 | 0.00 | (0.02) | -0.05 — 0.04 |
| **School-Reported Behavioral Outcome** | | | | | | | | | |
| ***Truancy—Unexcused Absence Rate (Percent)*[a]** | -0.38 | (0.26) | -0.89 — 0.12 | -0.65* | (0.31) | -1.25 — -0.05 | 0.26 | (0.40) | -0.52 — 1.05 |
| ***Misconduct*** | | | | | | | | | |
| Percent committing any infraction | 0.15 | (3.01) | -5.75 — 6.05 | 1.76 | (2.36) | -2.86 — 6.38 | -1.61 | (3.82) | -9.11 — 5.88 |
| Percent committing repeated infractions (2+) | -2.13 | (2.58) | -7.19 — 2.94 | -1.67 | (1.98) | -5.56 — 2.21 | -0.46 | (3.26) | -6.84 — 5.93 |
| ***Delinquency*** | | | | | | | | | |
| Percent committing any infraction | 1.77 | (3.66) | -5.41 — 8.94 | -4.24 | (2.33) | -8.80 — 0.33 | 6.00 | (3.77) | -1.39 — 13.40 |
| Percent committing repeated infractions (2+) | -2.31 | (3.04) | -8.28 — 3.65 | -1.75 | (1.73) | -5.14 — 1.64 | -0.57 | (2.77) | -6.00 — 4.87 |

[a]   Based on 27 sites that reported unexcused absences and total days enrolled.

*   p-value (of adjusted difference in means) < 0.05, two-tailed test.
+   p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit F.16: Subgroup Findings by Self-Reported Baseline Delinquency: Interpersonal Relationships, Personal Responsibility, and Community Involvement**

| Self-Reported Outcome | Any Delinquency | | | No Delinquency | | | Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| | Estimated Impact | Standard Error | 95% Confidence Interval | Estimated Impact | Standard Error | 95% Confidence Interval | Estimated Impact | Standard Error | 95% Confidence Interval |
| Pro-social Behaviors | -0.02 | 0.04 | -0.10 — 0.06 | -0.01 | 0.02 | -0.05 — 0.04 | -0.01 | (0.05) | -0.10 — 0.08 |

\* p-value (of adjusted difference in means) < 0.05, two-tailed test.

\+ p-value (of adjusted difference in means) < BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006and Spring 2007.

**Exhibit F.17: Subgroup Findings by Self-Reported Baseline Delinquency: Academic Outcomes**

| Self-Reported Outcome | Any Delinquency | | | No Delinquency | | | Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| | Estimated Impact | Standard Error | 95% Confidence Interval | Estimated Impact | Standard Error | 95% Confidence Interval | Estimated Impact | Standard Error | 95% Confidence Interval |
| Scholastic Efficacy and School Bonding | 0.07 | (0.04) | -0.02 — 0.15 | 0.03 | (0.02) | -0.02 — 0.07 | 0.04 | (0.05) | -0.06 — 0.14 |
| Future Orientation | 0.05 | (0.04) | -0.03 — 0.13 | 0.02 | (0.02) | -0.01 — 0.05 | 0.03 | (0.04) | -0.05 — 0.12 |
| **School-Reported Outcome** | | | | | | | | | |
| ***Overall Absenteeism Rate (Percent)*** | -0.13 | (0.56) | -1.22 — 0.96 | -0.61* | (0.25) | -1.09 — -0.13 | 0.49 | (0.61) | -0.71 — 1.68 |
| *Grades (Range 1–5)* | | | | | | | | | |
| Math | 0.00 | (0.09) | -0.19 — 0.18 | -0.08 | (0.05) | -0.18 — 0.02 | 0.07 | (0.11) | -0.14 — 0.28 |
| English Language Arts | 0.06 | (0.09) | -0.12 — 0.24 | -0.08 | (0.05) | -0.18 — 0.03 | 0.14 | (0.11) | -0.07 — 0.35 |
| Science | 0.03 | (0.09) | -0.16 — 0.21 | -0.06 | (0.05) | -0.16 — 0.04 | 0.09 | (0.11) | -0.12 — 0.30 |
| Social Studies | 0.09 | (0.10) | -0.10 — 0.28 | -0.07 | (0.05) | -0.17 — 0.04 | 0.16 | (0.11) | -0.06 — 0.38 |
| *State Assessment Tests* | | | | | | | | | |
| Math—Percent Proficient | 2.38 | (3.91) | -5.29 — 10.04 | -3.17 | (2.13) | -7.34 — 1.00 | 5.55 | (4.45) | -3.18 — 14.27 |
| Reading/ELA—Percent Proficient | -0.03 | (4.02) | -7.90 — 7.85 | -2.19 | (2.14) | -6.38 — 2.00 | 2.17 | (4.55) | -6.76 — 11.09 |

\*   p-value (of adjusted difference in means) < 0.05, two-tailed test.

+   p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➔ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing.

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

**Exhibit F.18: Subgroup Findings by Self-Reported Baseline Delinquency: Delinquent Behaviors and Participation in Harmful Activities**

| Self-Reported Outcome | Any Delinquency | | | No Delinquency | | | Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| | Estimated Impact | Standard Error | 95% Confidence Interval | Estimated Impact | Standard Error | 95% Confidence Interval | Estimated Impact | Standard Error | 95% Confidence Interval |
| Misconduct | 0.02 | (0.05) | -0.07 — 0.11 | -0.01 | (0.02) | -0.06 — 0.03 | 0.03 | (0.05) | -0.07 — 0.13 |
| Delinquency | 0.03 | (0.03) | -0.03 — 0.08 | 0.00 | (0.01) | -0.01 — 0.02 | 0.02 | (0.03) | -0.03 — 0.08 |
| **School-Reported Behavioral Outcome** | | | | | | | | | |
| ***Truancy—Unexcused Absence Rate (Percent)*[a]** | -0.53 | (0.46) | -1.44 — 0.38 | -0.41 | (0.22) | -0.84 — 0.02 | -0.12 | (0.51) | -1.12 — 0.88 |
| ***Misconduct*** | | | | | | | | | |
| Percent committing any infraction | 4.59 | (3.67) | -2.60 — 11.78 | 1.36 | (1.88) | -2.33 — 5.06 | 3.22 | (4.12) | -4.86 — 11.31 |
| Percent committing repeated infractions (2+) | 2.72 | (3.21) | -3.57 — 9.01 | -2.32 | (1.53) | -5.32 — 0.67 | 5.04 | (3.55) | -1.92 — 12.01 |
| ***Delinquency*** | | | | | | | | | |
| Percent committing any infraction | -4.36 | (3.66) | -11.53 — 2.82 | -0.36 | (1.79) | -3.87 — 3.14 | -3.99 | (4.08) | -11.99 — 4.00 |
| Percent committing repeated infractions (2+) | -0.53 | (3.04) | -6.49 — 5.43 | -0.15 | (1.26) | -2.61 — 2.32 | -0.38 | (3.29) | -6.84 — 6.07 |

[a]    Based on 27 sites that reported unexcused absences and total days enrolled.

\*    p-value (of adjusted difference in means) < 0.05, two-tailed test.

+    p-value (of adjusted difference in means) <  BH-Corrected Critical Value ➜ statistically significant at the 0.05 level correcting for the false discovery rate under multiple testing

*Source:* Impact Evaluation of the U.S. Department of Education's Student Mentoring Program—Student Survey, Fall 2005 and Spring 2006; Fall 2006 and Spring 2007; School Records, SY 2004-2005, SY 2005-2006, SY 2006-2007.

# Appendix G: Site-Level Predictors and Impacts

**Exhibit G.1: Site-Level Impact Estimates on Parental Relationships, Personal Responsibility, and Community Involvement with 95 Percent Confidence Intervals**
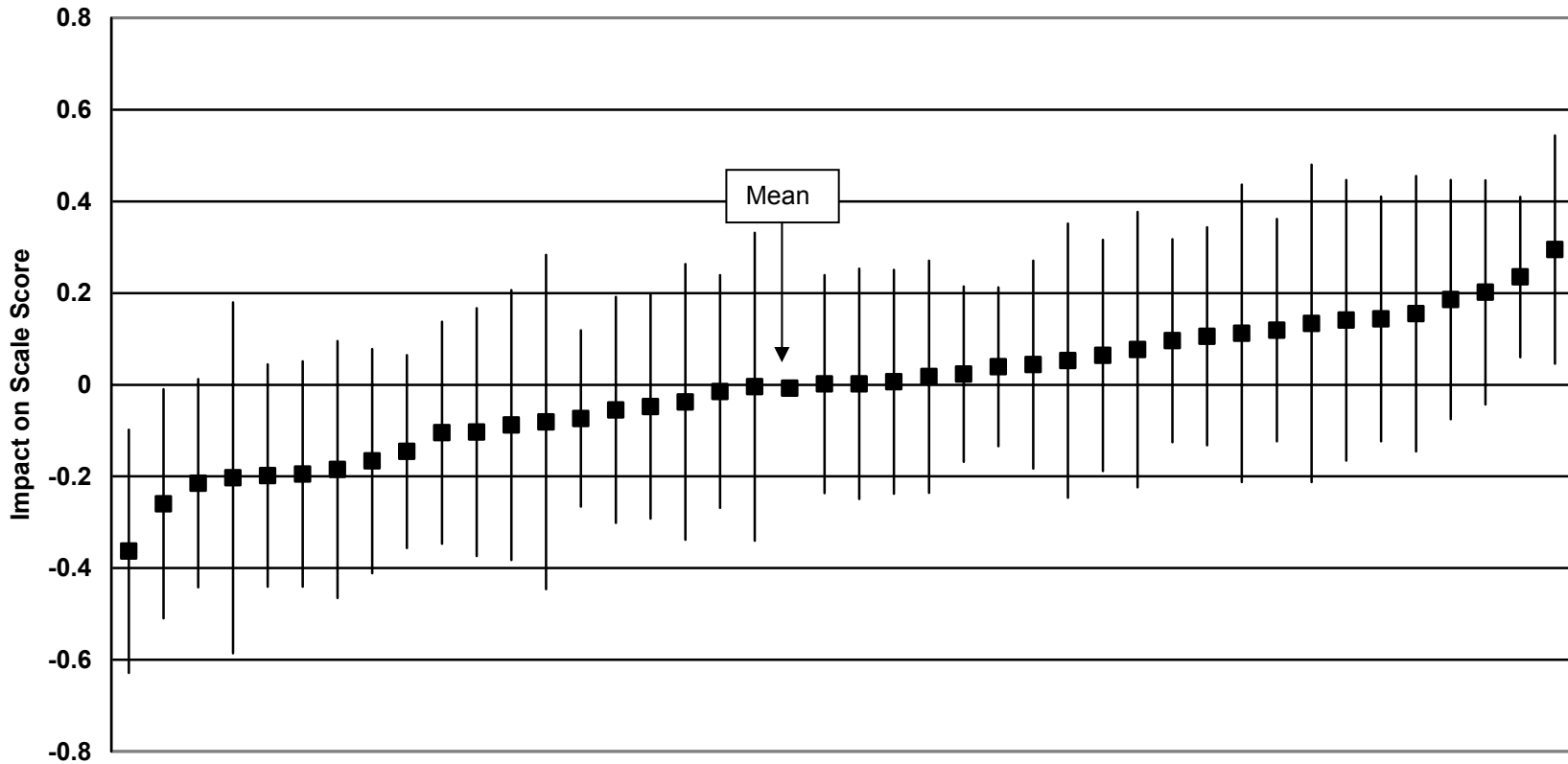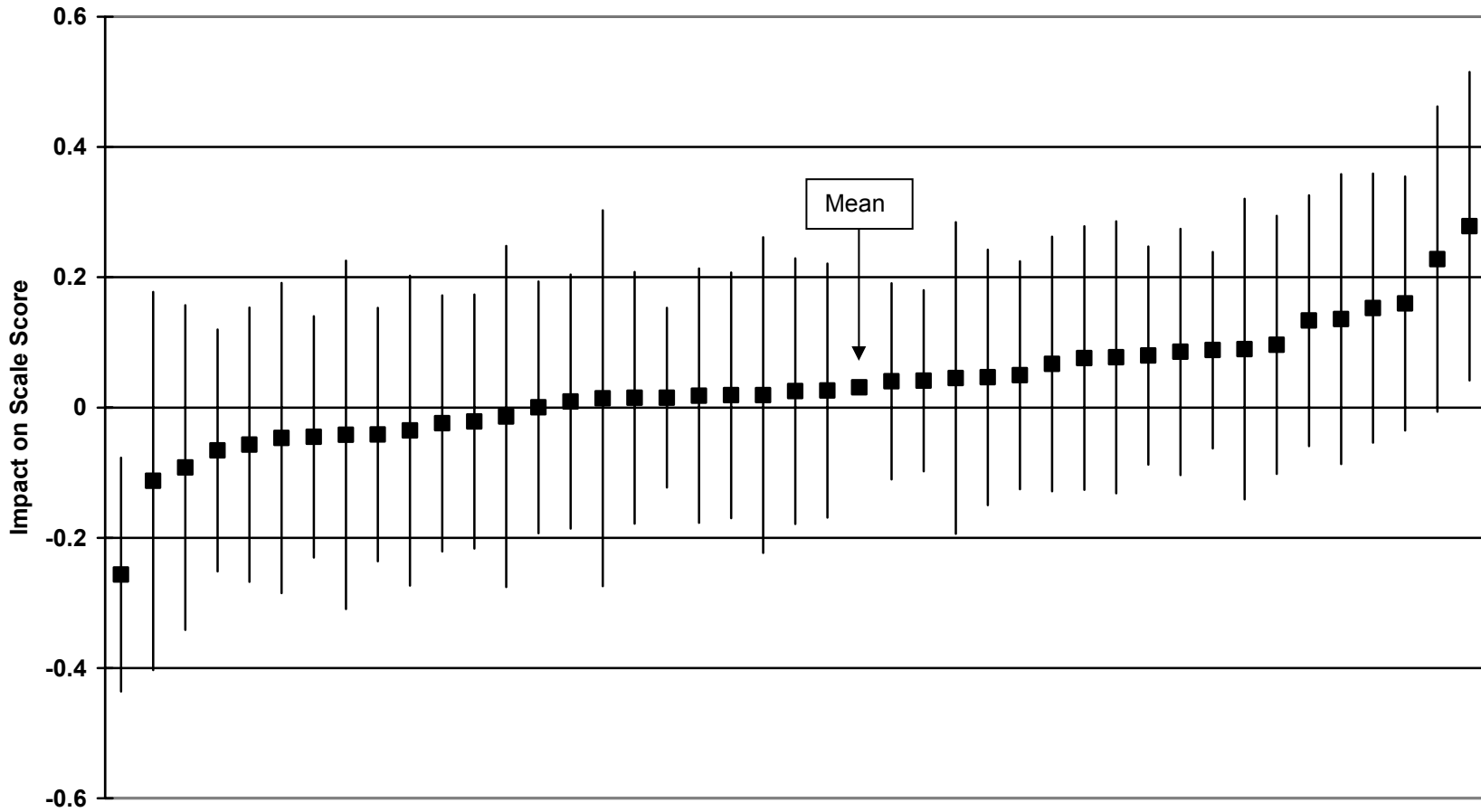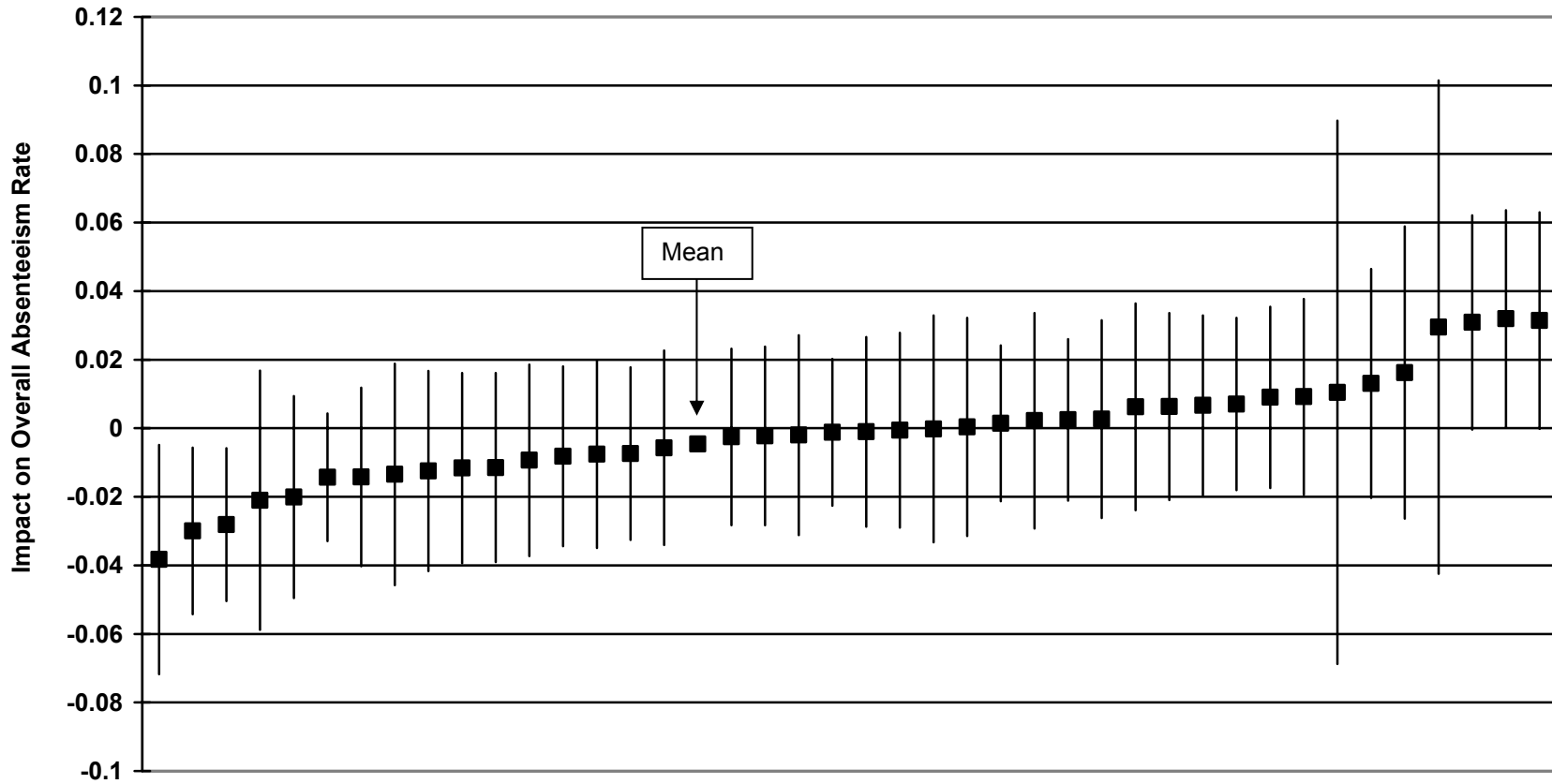


Notes: Site-level impact estimates based on regression-adjusted differences between treatment and control group means.  Site-level Ns range from 25 to 115.
Source: Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – Student Survey, Spring 2006 and Spring 2007
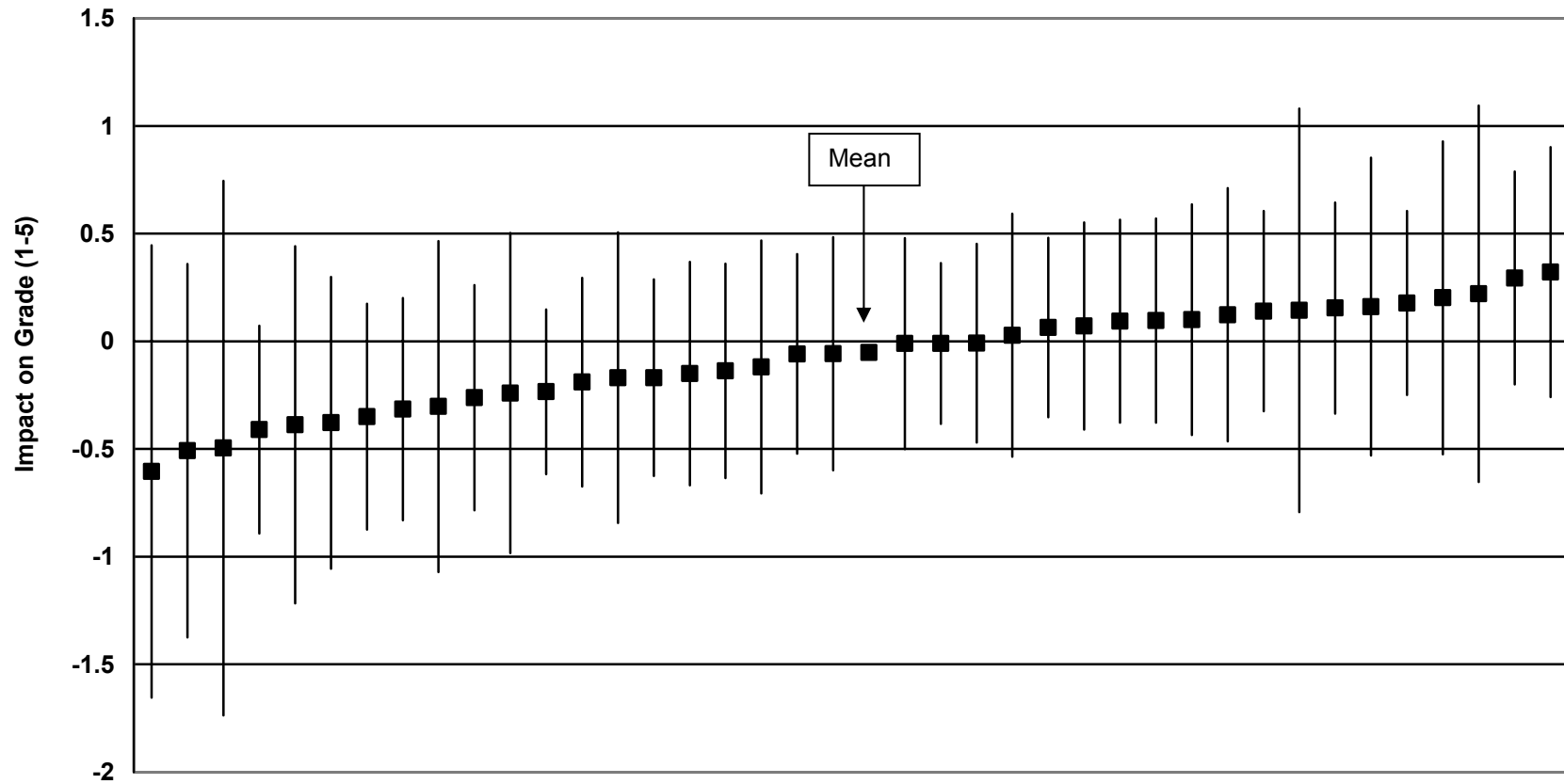
**Exhibit G.2: Site-Level Impact Estimates on Scholastic Efficacy and School Bonding with 95 Percent Confidence Intervals**
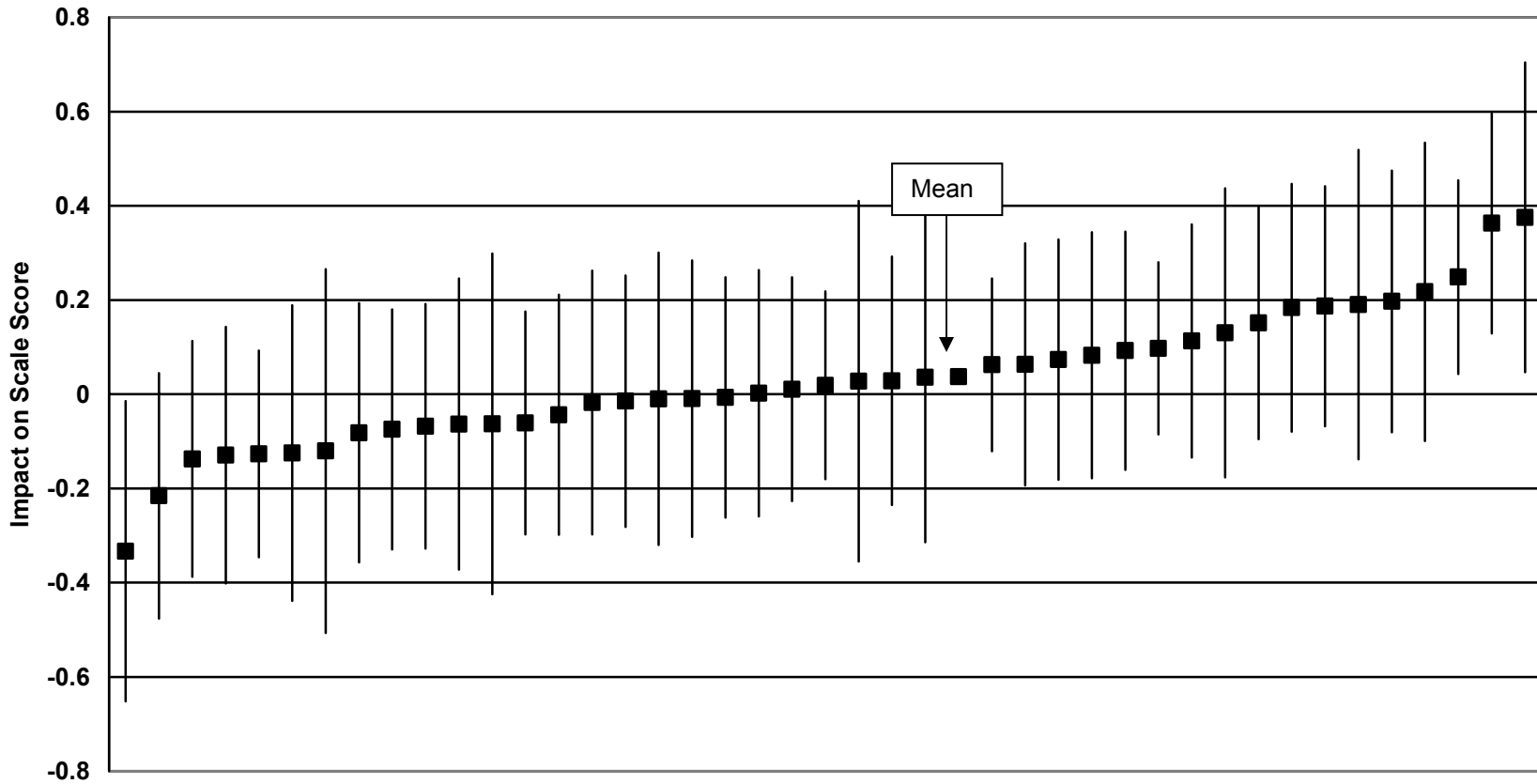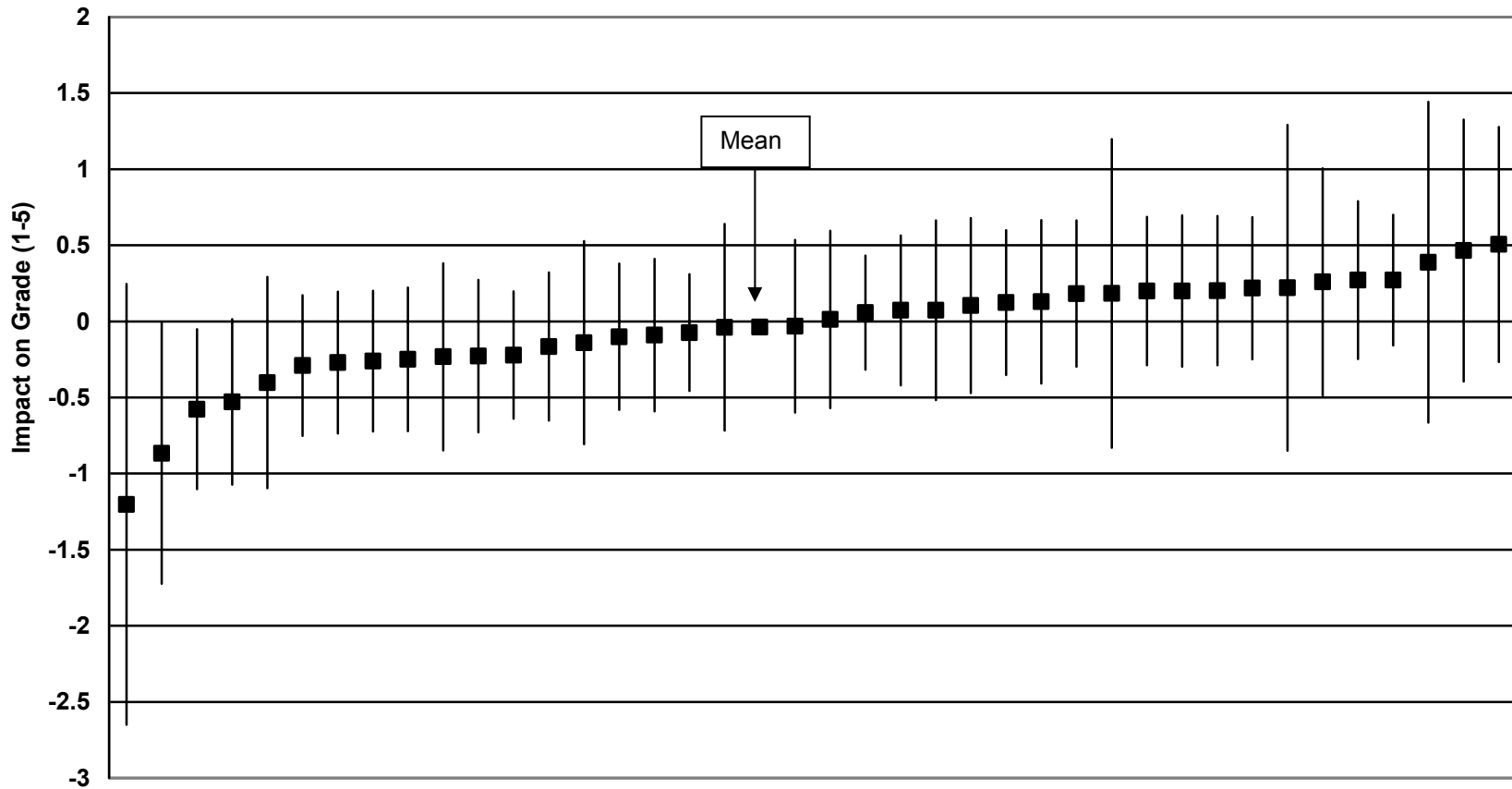


Notes: Site-level impact estimates based on regression-adjusted differences between treatment and control group means. Site-level Ns range from 25 to 114.
Source: Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – Student Survey, Spring 2006 and Spring 2007

**Exhibit G.3: Site-Level Impact Estimates on Future Orientation with 95 Percent Confidence Intervals**



Notes: Site-level impact estimates based on regression-adjusted differences between treatment and control group means.  Site-level Ns range from 26 to 114.

Source: Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – Student Survey, Spring 2006 and Spring 2007

**Exhibit G.4: Site-Level Impact Estimates on Overall Absenteeism Rate with 95 Percent Confidence Intervals**



Notes: Site-level impact estimates based on regression-adjusted differences between treatment and control group means.  Site-level Ns range from 9 to 122.
Source: Impact Evaluation of the U.S. Department of Education's Student Mentoring Program –School Records, Spring 2006 and Spring 2007
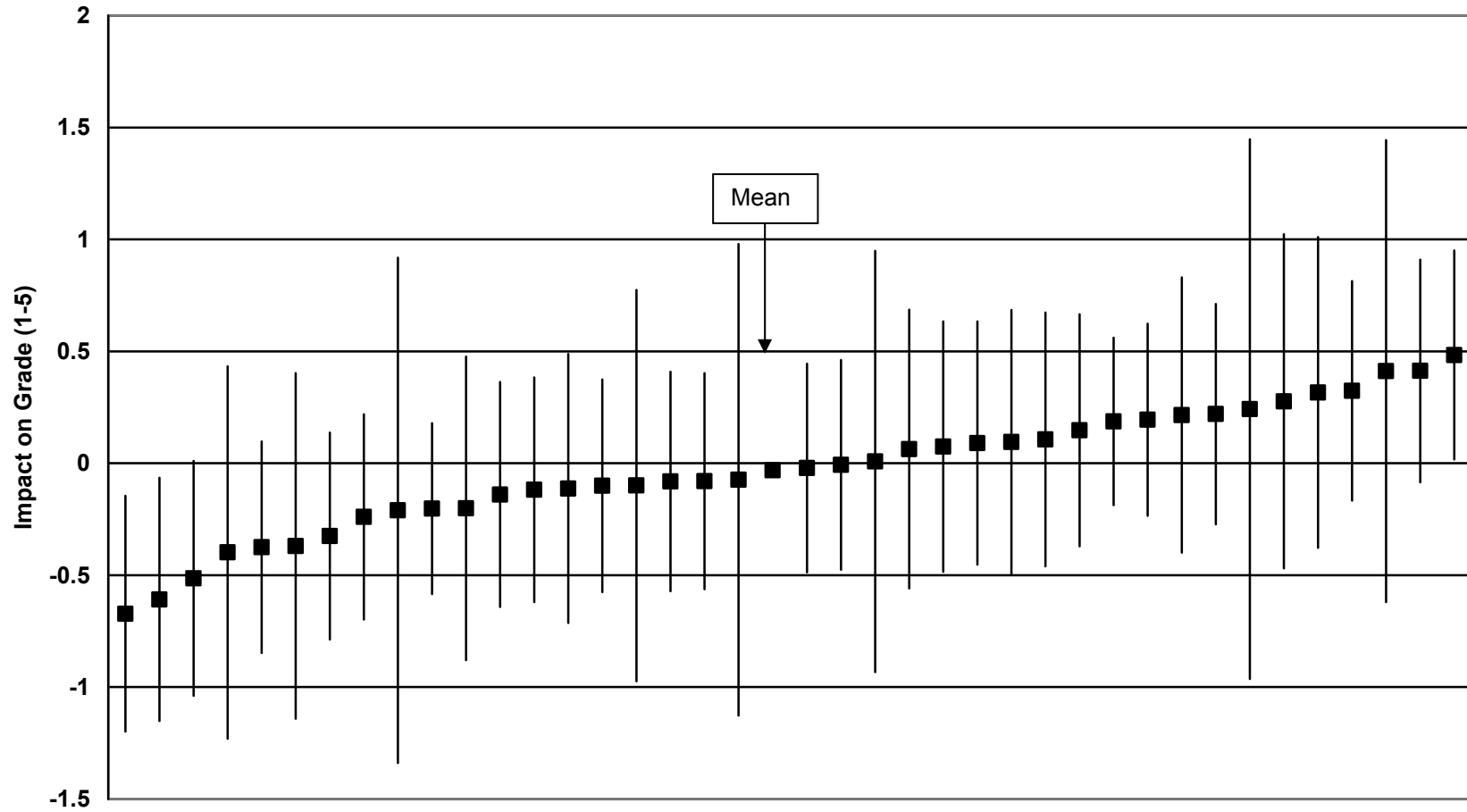
**Exhibit G.5: Site-Level Impact Estimates on Math Grades with 95 Percent Confidence Intervals**
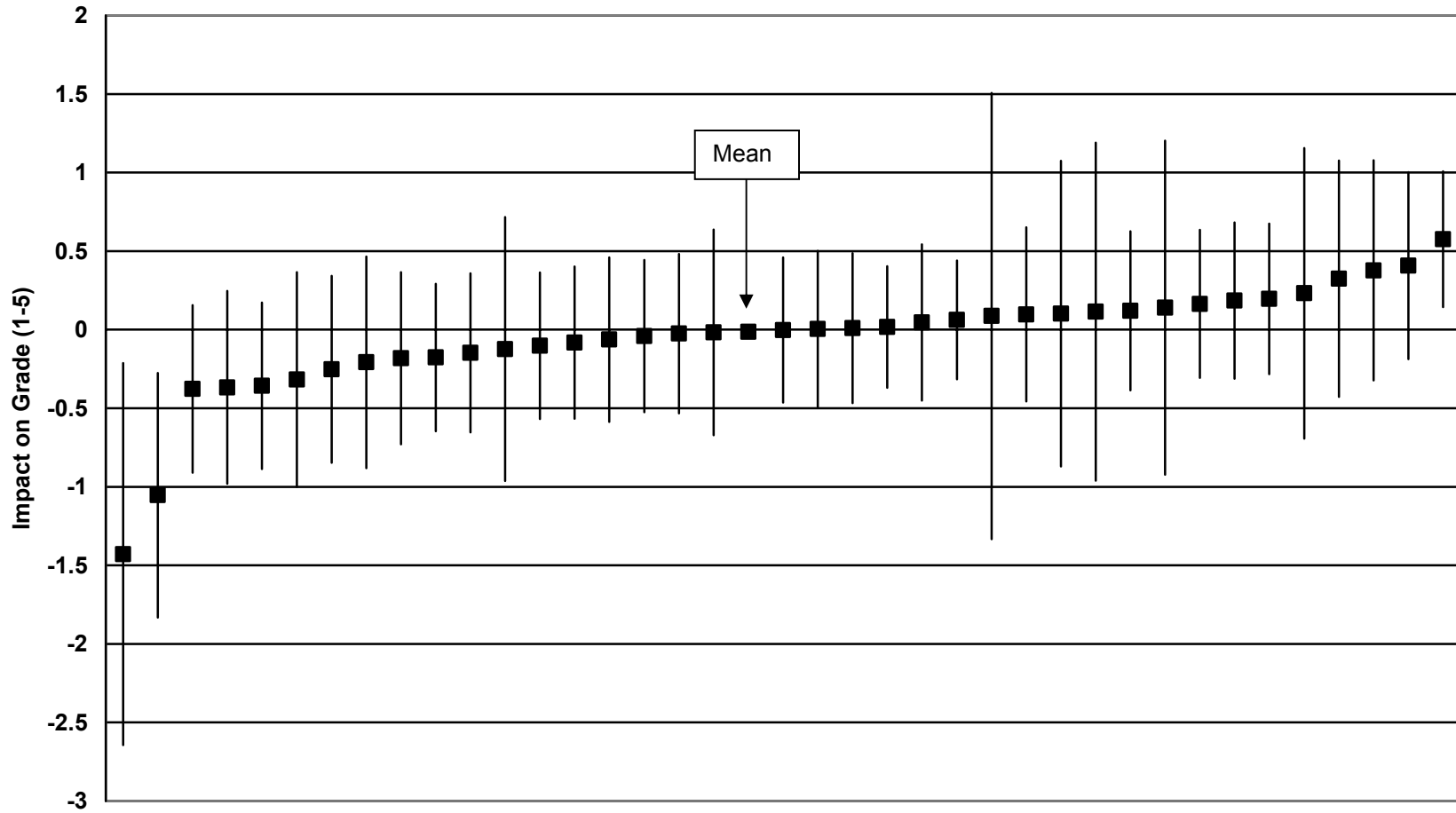


Notes: Site-level impact estimates based on regression-adjusted differences between treatment and control group means.  Site-level Ns range from 17 to 93.
Source: Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – School Records, Spring 2006 and Spring 2007

**Exhibit G.6: Site-Level Impact Estimates on English Language Arts Grades with 95 Percent Confidence Intervals**

Notes: Site-level impact estimates based on regression-adjusted differences between treatment and control group means.  Site-level Ns range from 16 to 93.
Source: Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – School Records, Spring 2006 and Spring 2007

**Exhibit G.7:  Site-Level Impact Estimates on Science Grades with 95 Percent Confidence Intervals**
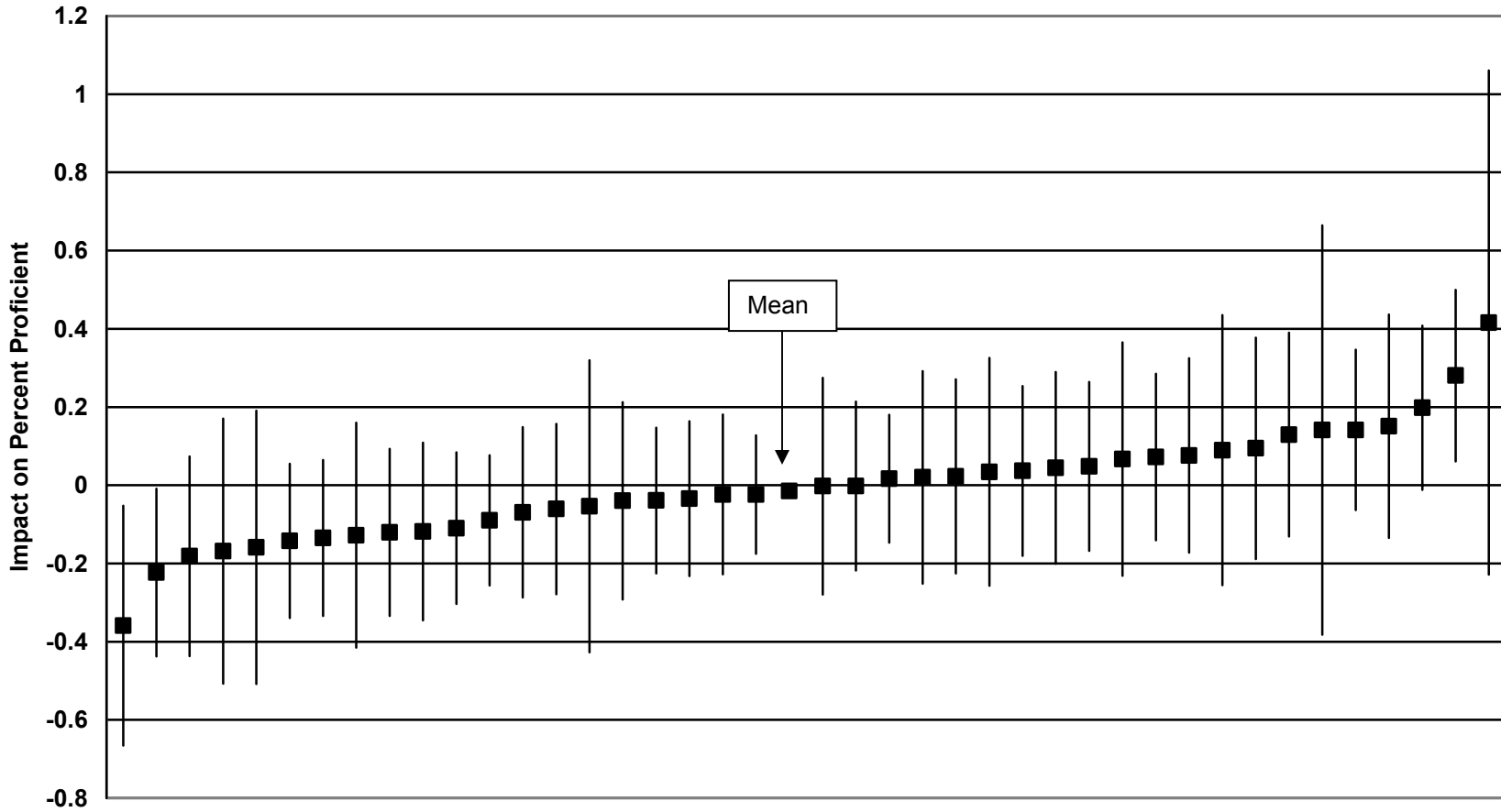


Notes: Site-level impact estimates based on regression-adjusted differences between treatment and control group means.  Site-level Ns range from 13 to 93.
Source: Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – School Records, Spring 2006 and Spring 2007

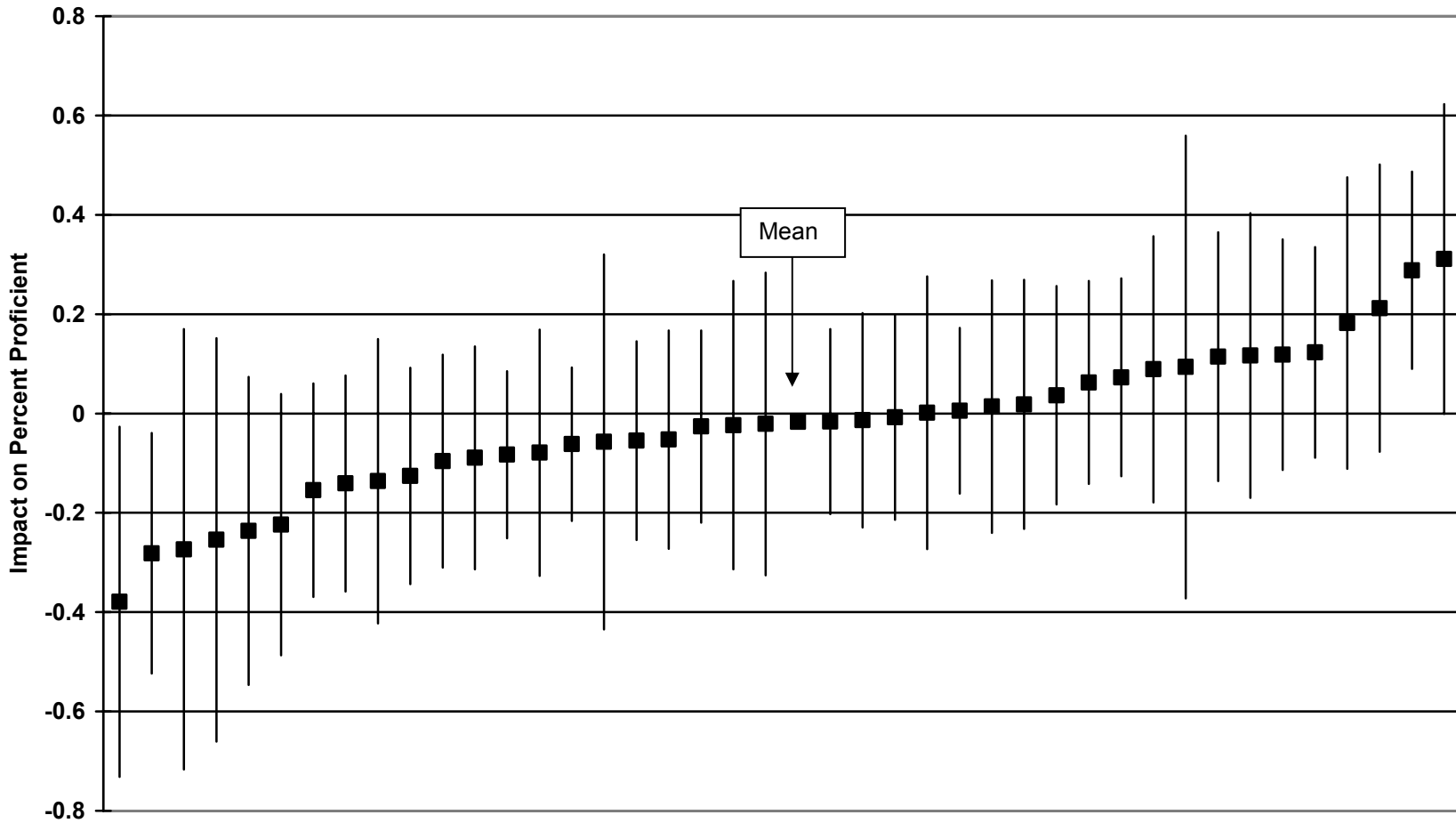**Exhibit G.8: Site-Level Impact Estimates on Social Studies Grades with 95 Percent Confidence Intervals**



Notes: Site-level impact estimates based on regression-adjusted differences between treatment and control group means.  Site-level Ns range from 13 to 93.
Source: Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – School Records, Spring 2006 and Spring 2007

**Exhibit G.9: Site-Level Impact Estimates on Percent Students Proficient in Math with 95 Percent Confidence Intervals**



Notes: Site-level impact estimates based on regression-adjusted differences between treatment and control group means.  Site-level Ns range from 16 to 115.

Source: Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – School Records, Spring 2006 and Spring 2007

**Exhibit G.10: Site-Level Impact Estimates on Percent Students Proficient in Reading/English Language Arts (ELA) with 95 Percent Confidence Intervals**
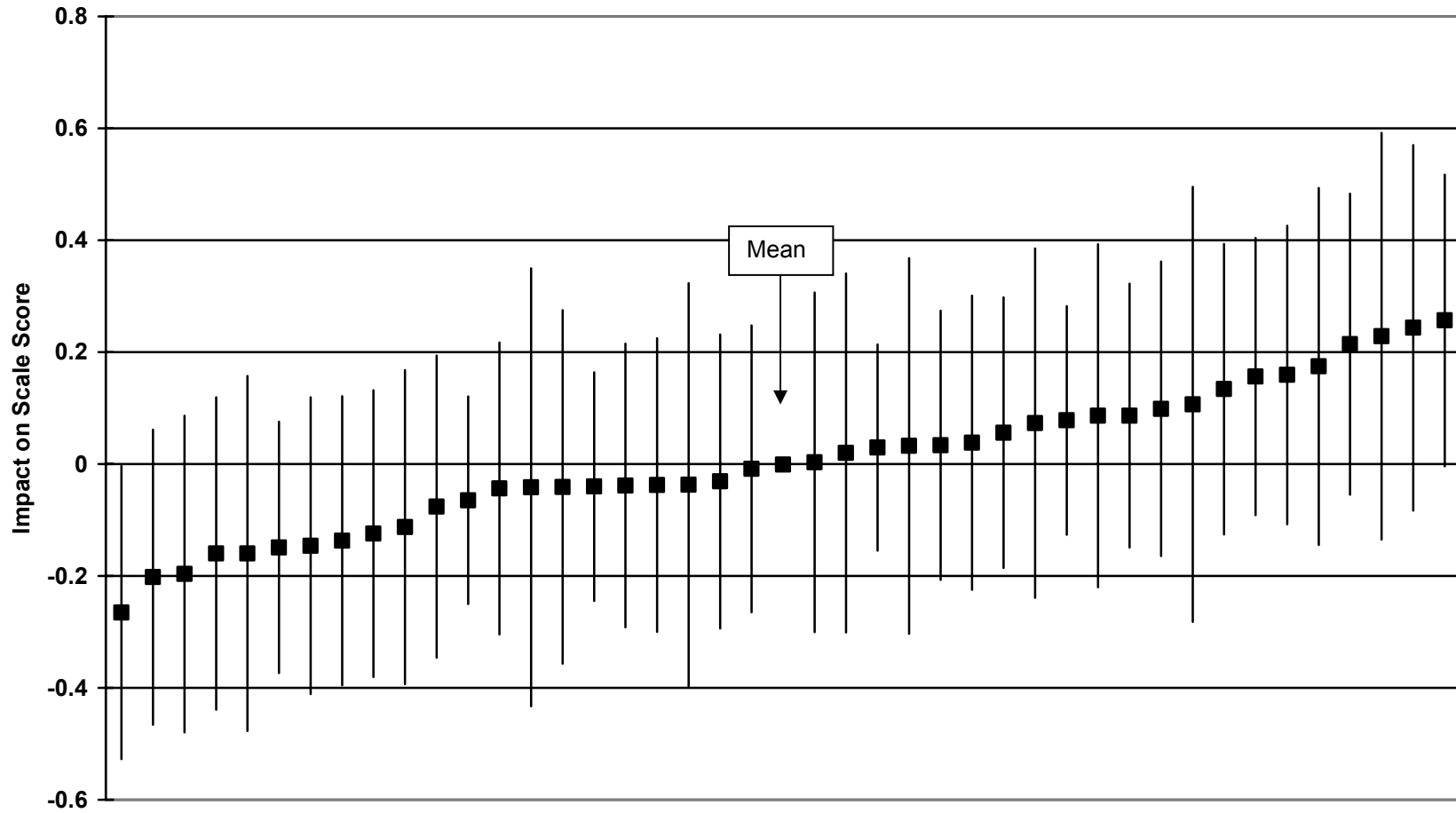


Notes: Site-level impact estimates based on regression-adjusted differences between treatment and control group means.  Site-level Ns range from 16 to 114.

Source: Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – School Records, Spring 2006 and Spring 2007

**Exhibit G.11: Site-Level Impact Estimates on Self-Reported Misconduct with 95 Percent Confidence Intervals**
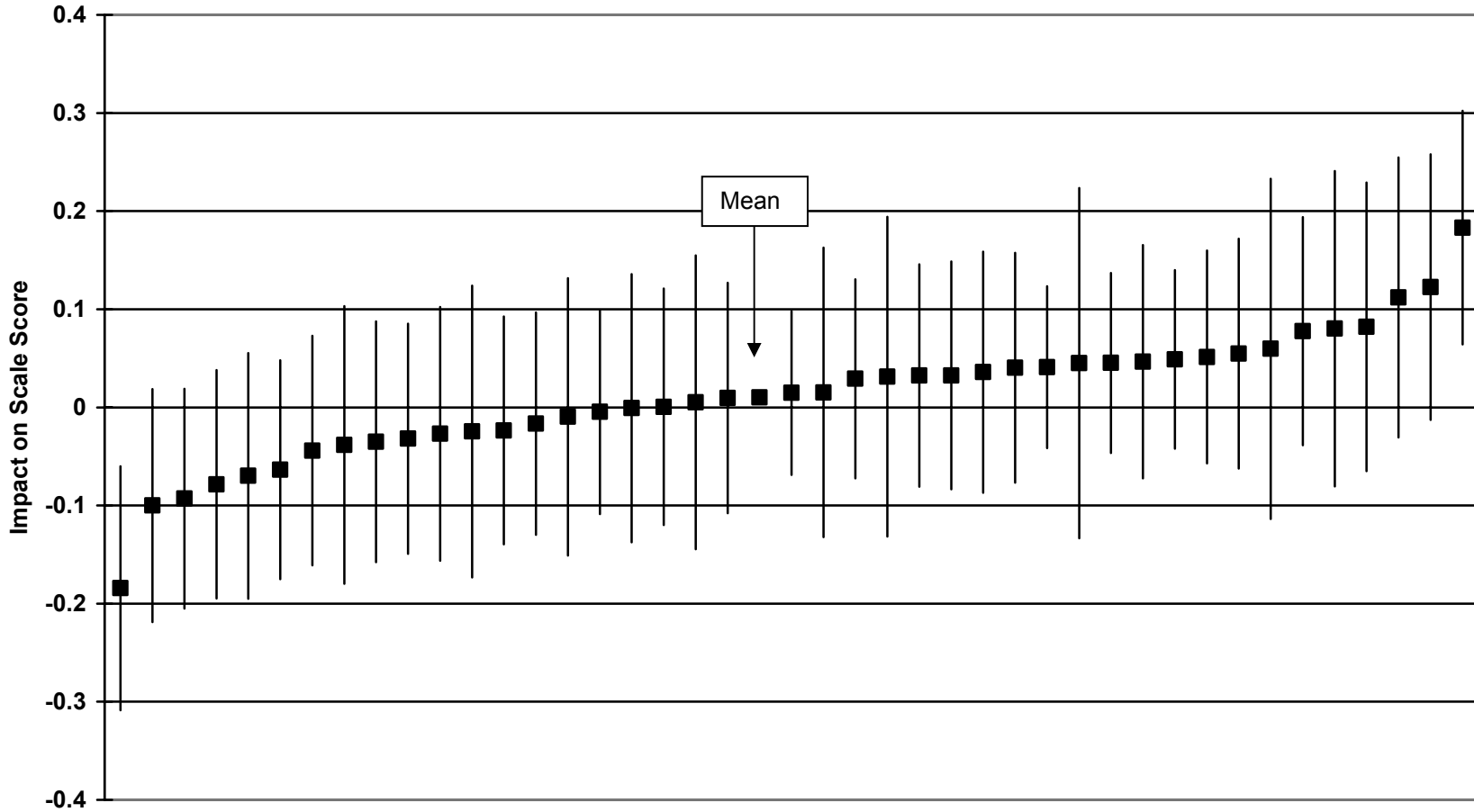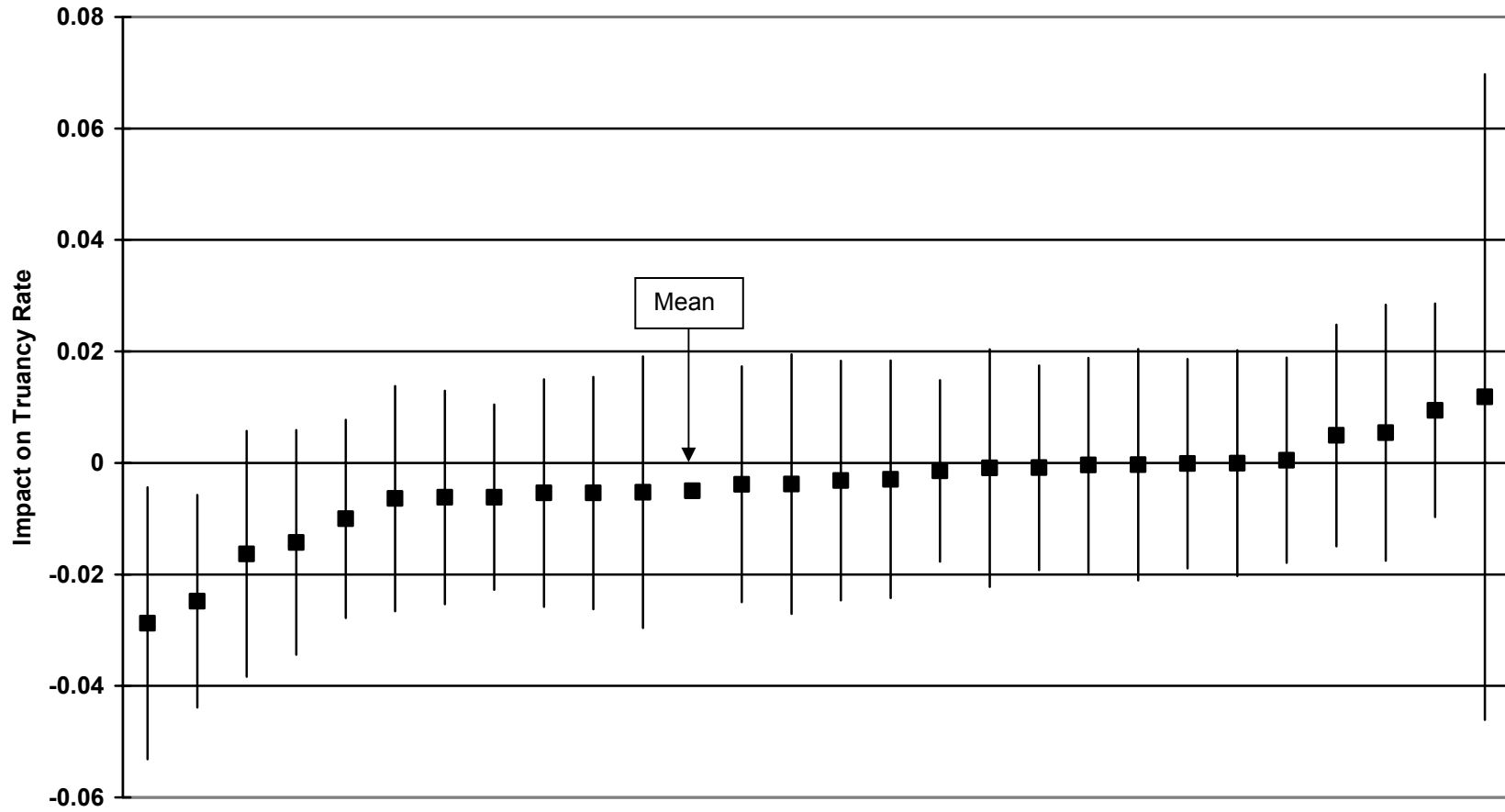


Notes: Site-level impact estimates based on regression-adjusted differences between treatment and control group means. Site-level Ns range from 26 to 116.
Source: Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – Student Survey, Spring 2006 and Spring 2007

**Exhibit G.12: Site-Level Impact Estimates on Self-Reported Delinquency with 95 Percent Confidence Intervals**

Notes: Site-level impact estimates based on regression-adjusted differences between treatment and control group means.  Site-level Ns range from 26 to 115.
Source: Impact Evaluation of the U.S. Department of Education's Student Mentoring Program –Student Survey, Spring 2006 and Spring 2007

**Exhibit G.13: Site-Level Impact Estimates on Truancy Rate with 95 Percent Confidence Intervals**



Notes: Site-level impact estimates based on regression-adjusted differences between treatment and control group means. Site-level Ns range from 9 to 85.
Source: Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – School Records, Spring 2006 and Spring 2007

**Exhibit G.14: Site-Level Impact Estimates on School-Reported Misconduct, Any Infraction with 95 Percent Confidence Intervals**
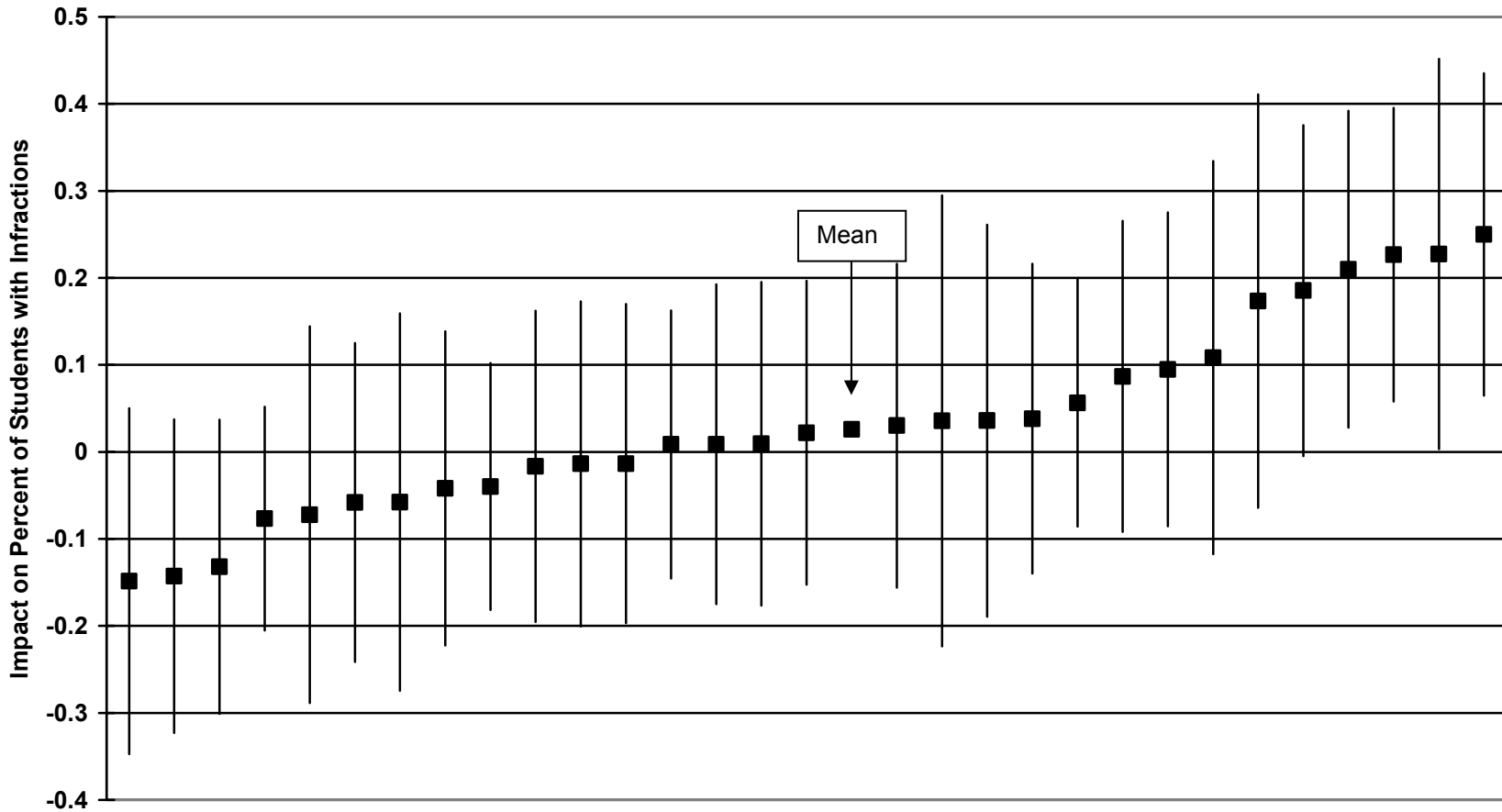


Notes: Site-level impact estimates based on regression-adjusted differences between treatment and control group means.  Site-level Ns range from 30 to 122.
Source: Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – School Records, Spring 2006 and Spring 2007

**Exhibit G.15: Site-Level Impact Estimates on School-Reported Misconduct—Repeated Infractions with 95 Percent Confidence Intervals**
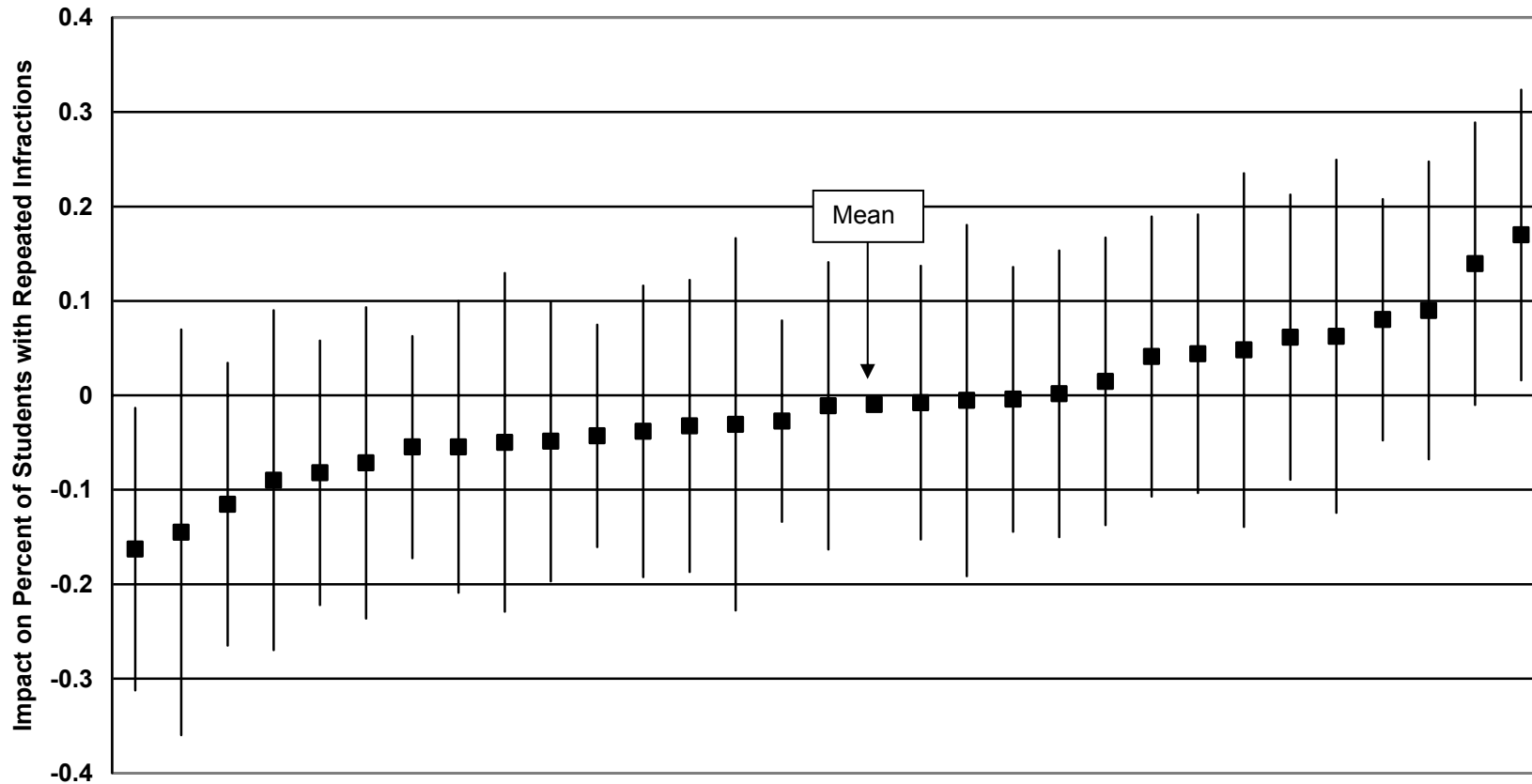


Notes: Site-level impact estimates based on regression-adjusted differences between treatment and control group means.  Site-level Ns range from 30 to 122.

Source: Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – School Records, Spring 2006 and Spring 2007

**Exhibit G.16: Site-Level Impact Estimates on School-Reported Delinquency—Any Infraction with 95 Percent Confidence Intervals**



Notes: Site-level impact estimates based on regression-adjusted differences between treatment and control group means. Site-level Ns range from 30 to 122.
Source: Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – School Records, Spring 2006 and Spring 2007

**Exhibit G.17: Site-Level Impact Estimates on School-Reported Delinquency—Repeated Infractions with 95 Percent Confidence Intervals**
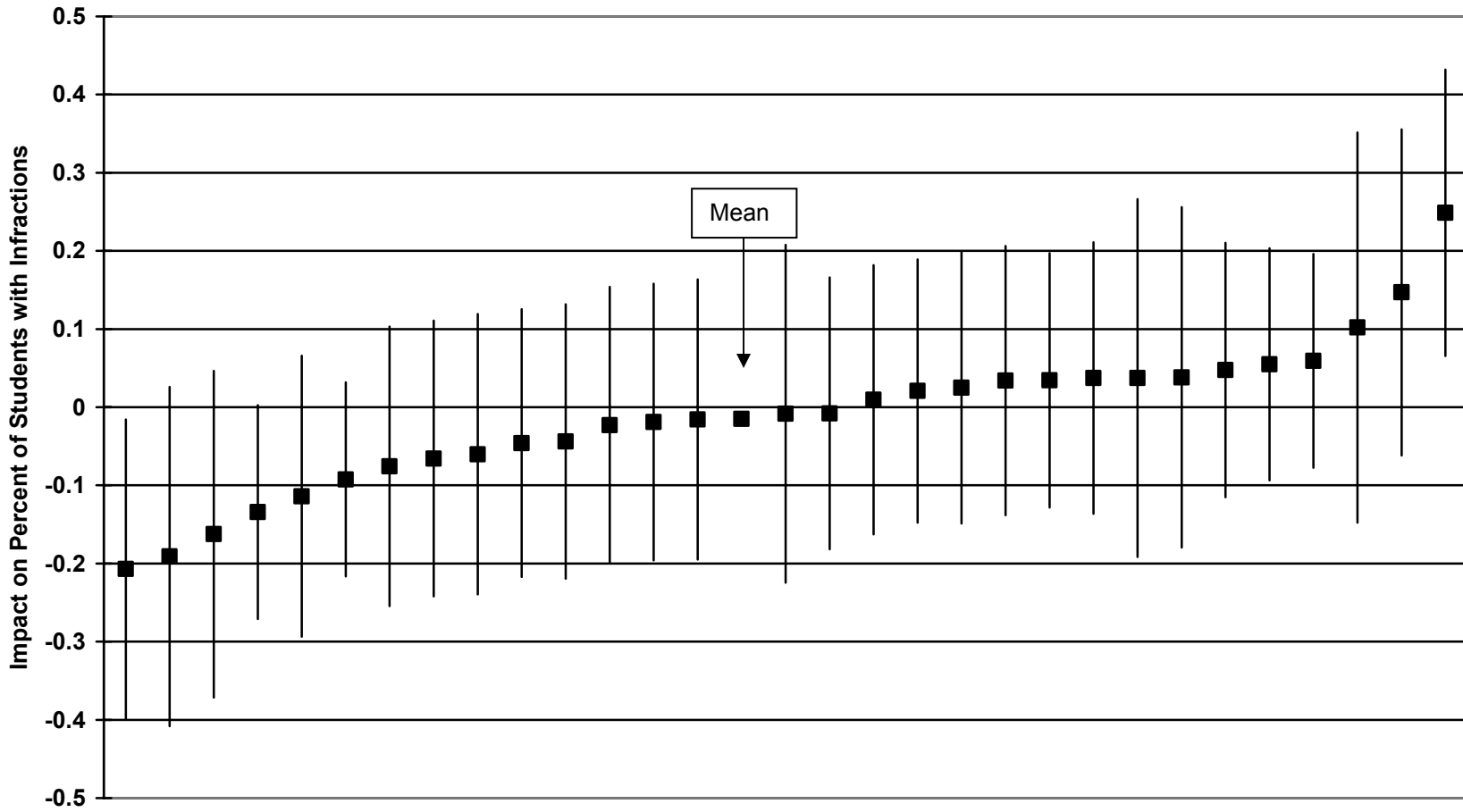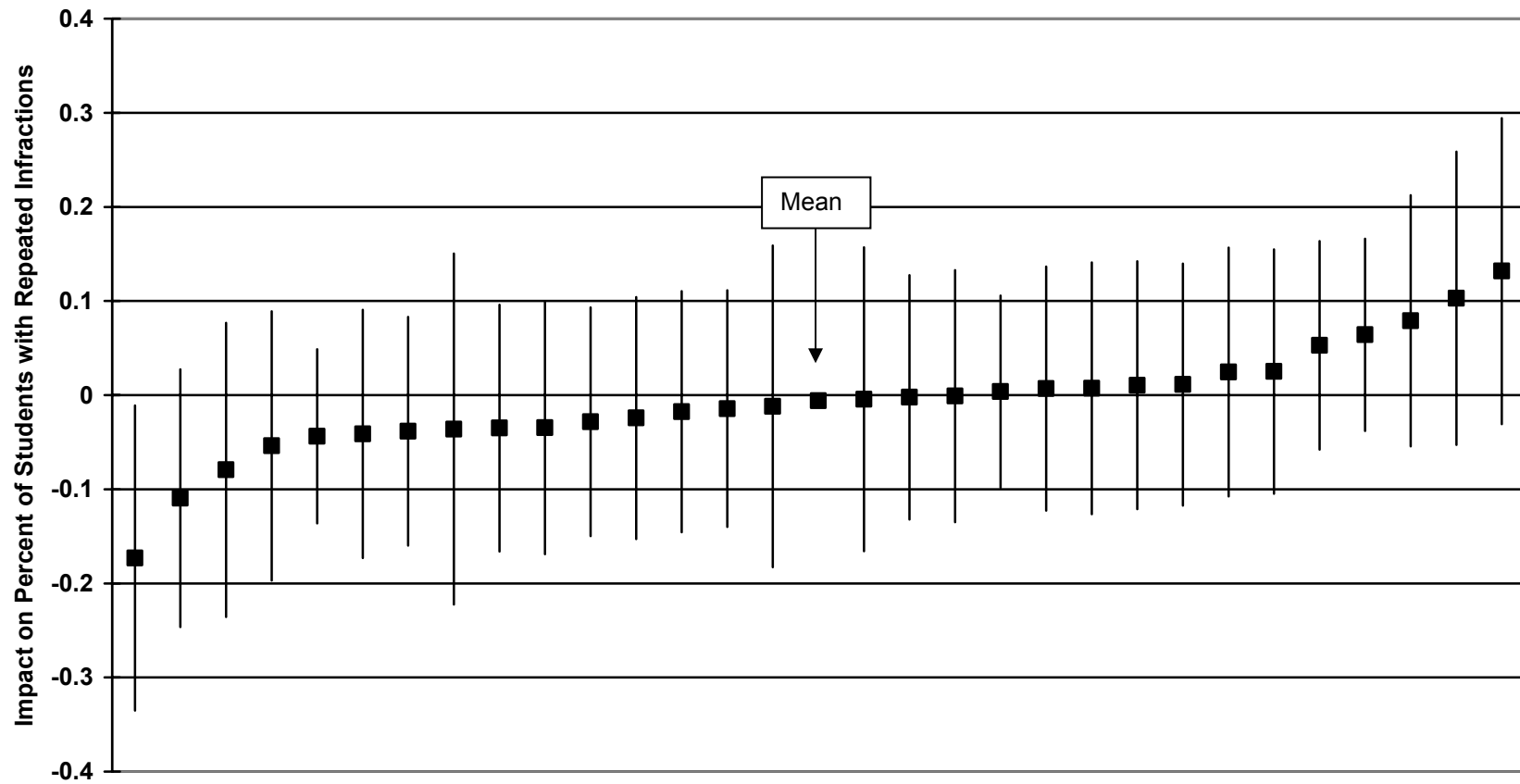


Notes: Site-level impact estimates based on regression-adjusted differences between treatment and control group means.  Site-level Ns range from 30 to 122.
Source: Impact Evaluation of the U.S. Department of Education's Student Mentoring Program – School Records, Spring 2006 and Spring 2007

**Exhibit G.18**

**Site-Level Associations: Relationship between Program Characteristics and Pro-social Behaviors**

| | Bivariate Specifications (p-values) | Multivariate Specification (p-values) |
|---|---|---|
| *Program Delivery (based on pre-intervention activities or characteristics of mentors)* | | |
| Average hours of mentor pre-match training provided to mentors | -0.01 | 0.00 |
| | (0.33) | (0.86) |
| Percent of mentors 22 years or below | 0.13 | 0.06 |
| | (0.08) | (0.48) |
| Percent of mentor/student matches of the same race/ethnicity | 0.09 | 0.07 |
| | (0.31) | (0.54) |
| *Program Delivery (based on aggregated mentor reports post-intervention)* | | |
| Amount of ongoing mentor support (average frequency of mentor-supervisor meetings) | -0.08* | -0.08* |
| | (0.01) | (0.02) |
| Frequency of working on relationship building in student-mentor meetings | 0.07 | 0.16 |
| | (0.63) | (0.33) |
| Percent of mentor/student matches lasting 6 months or longer | -0.06 | -0.03 |
| | (0.53) | (0.76) |
| Average total hours of mentor/student meetings per month | 0.00 | 0.01 |
| | (0.50) | (0.28) |
| *Student Characteristics* | | |
| Percent of students with self-reported delinquent behaviors at baseline | 0.03 | -0.12 |
| | (0.90) | (0.61) |
| Percent of students scoring "not proficient" in either math or reading/ELA at baseline | -0.10 | -0.06 |
| | (0.31) | (0.62) |
| *Percent of Control Group Students Receiving Mentoring* | 0.08 | 0.14 |
| | (0.72) | (0.51) |

\* p-value<.05, two-tailed test

**Exhibit G.19**

**Site-Level Associations: Relationship between Program Characteristics and Academic Outcomes**

| | Scholastic Efficacy & School Bonding | | Future Orientation | | Overall Absenteeism Rate | | Math Grades | | English Language Arts Grades | | Science Grades | | Social Studies Grades | | Math Proficiency | | Reading/ELA Proficiency | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Bivariate Specifications (p-value) | Multivariate Specification (p-value) | Bivariate Specifications (p-value) | Multivariate Specification (p-value) | Bivariate Specifications (p-value) | Multivariate Specification (p-value) | Bivariate Specifications (p-value) | Multivariate Specification (p-value) | Bivariate Specifications (p-value) | Multivariate Specification (p-value) | Bivariate Specifications (p-value) | Multivariate Specification (p-value) | Bivariate Specifications (p-value) | Multivariate Specification (p-value) | Bivariate Specifications (p-value) | Multivariate Specification (p-value) | Bivariate Specifications (p-value) | Multivariate Specification (p-value) |
| *Program Delivery (based on pre-intervention activities or characteristics of mentors)* | | | | | | | | | | | | | | | | | | |
| Average hours of mentor pre-match training provided to mentors | -0.01 (0.46) | 0.00 (0.74) | -0.01 (0.30) | -0.01 (0.30) | 0.00 (0.17) | 0.00 (0.15) | -0.01 (0.42) | -0.03 (0.06) | 0.01 (0.73) | 0.01 (0.75) | 0.00 (0.98) | -0.01 (0.79) | -0.02 (0.48) | -0.02 (0.40) | 0.00 (0.93) | 0.00 (0.71) | -0.01 (0.58) | -0.01 (0.29) |
| Percent of mentors 22 years or below | 0.05 (0.47) | 0.05 (0.57) | -0.02 (0.71) | -0.03 (0.62) | 0.00 (0.31) | 0.01 (0.10) | -0.19 (0.12) | -0.29 (0.04) | 0.19 (0.29) | 0.23 (0.25) | -0.12 (0.40) | -0.13 (0.49) | 0.19 (0.31) | 0.08 (0.71) | 0.06 (0.36) | 0.06 (0.46) | 0.02 (0.83) | -0.05 (0.61) |
| Percent of mentor/student matches of the same race/ethnicity | -0.02 (0.82) | -0.01 (0.96) | 0.00 (0.99) | 0.02 (0.80) | 0.00 (0.97) | -0.01 (0.51) | 0.02 (0.88) | 0.29 (0.12) | 0.28 (0.21) | 0.56* (0.04) | -0.11 (0.55) | 0.18 (0.48) | -0.02 (0.92) | 0.06 (0.83) | -0.06 (0.51) | -0.03 (0.83) | 0.12 (0.21) | 0.12 (0.39) |
| *Program Delivery (based on aggregated mentor reports post-intervention)* | | | | | | | | | | | | | | | | | | |
| Amount of ongoing mentor support (average frequency of mentor-supervisor meetings) | -0.01 (0.72) | -0.01 (0.71) | -0.02 (0.23) | -0.03 (0.17) | 0.00 (0.95) | 0.01 (0.50) | -0.07 (0.15) | -0.11 (0.05) | -0.03 (0.64) | 0.03 (0.67) | -0.07 (0.22) | -0.06 (0.41) | -0.19* (0.01) | -0.22* (0.01) | -0.05 (0.06) | -0.05 (0.13) | -0.01 (0.70) | -0.02 (0.55) |
| Frequency of working on academic skills or homework in student-mentor meetings | -0.02 (0.89) | 0.00 (0.99) | -0.07 (0.43) | -0.07 (0.53) | -0.01 (0.49) | -0.01 (0.35) | -0.08 (0.73) | -0.03 (0.90) | -0.09 (0.79) | -0.27 (0.47) | -0.51 (0.07) | -0.45 (0.22) | -0.44 (0.24) | 0.00 (0.99) | -0.11 (0.42) | 0.07 (0.69) | 0.03 (0.84) | 0.01 (0.94) |
| Percent of mentor/student matches lasting 6 months or longer | 0.01 (0.91) | 0.11 (0.38) | 0.08 (0.18) | 0.13 (0.09) | -0.06* (0.01) | -0.04* (0.05) | 0.14 (0.36) | 0.09 (0.63) | -0.26 (0.25) | -0.26 (0.32) | 0.00 (0.99) | 0.12 (0.64) | -0.04 (0.87) | 0.30 (0.31) | -0.11 (0.21) | 0.02 (0.86) | -0.07 (0.45) | -0.08 (0.53) |
| Average total hours of mentor/student meetings per month | 0.01 (0.08) | 0.02 (0.08) | 0.01 (0.07) | 0.01* (0.04) | -0.01 (0.44) | -0.01 (0.48) | -0.03 (0.01) | -0.03* (0.02) | -0.05* (0.00) | -0.06* (0.00) | -0.01 (0.38) | -0.01 (0.59) | 0.00 (0.98) | -0.01 (0.55) | 0.00 (0.62) | -0.01 (0.40) | 0.00 (0.91) | -0.01 (0.45) |
| *Student Characteristics* | | | | | | | | | | | | | | | | | | |
| Percent of students with self-reported delinquent behaviors at baseline | -0.04 (0.86) | -0.15 (0.57) | 0.09 (0.50) | 0.04 (0.82) | 0.01 (0.31) | 0.00 (0.77) | -0.07 (0.85) | 0.22 (0.54) | -0.09 (0.86) | 0.09 (0.86) | -0.10 (0.80) | -0.07 (0.89) | 0.88 (0.08) | 1.39* (0.02) | 0.25 (0.21) | 0.37 (0.12) | 0.30 (0.15) | 0.32 (0.23) |

**Exhibit G.19**

**Site-Level Associations: Relationship between Program Characteristics and Academic Outcomes**

| | Scholastic Efficacy & School Bonding | | Future Orientation | | Overall Absenteeism Rate | | Math Grades | | English Language Arts Grades | | Science Grades | | Social Studies Grades | | Math Proficiency | | Reading/ELA Proficiency | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Bivariate Specifica-tions (p-value) | Multivariate Specifica-tion (p-value) | Bivariate Specifica-tions (p-value) | Multivariate Specifica-tion (p-value) | Bivariate Specifica-tions (p-value) | Multivariate Specifica-tion (p-value) | Bivariate Specifica-tions (p-value) | Multivariate Specifica-tion (p-value) | Bivariate Specifica-tions (p-value) | Multivariate Specifica-tion (p-value) | Bivariate Specifica-tions (p-value) | Multivariate Specifica-tion (p-value) | Bivariate Specifica-tions (p-value) | Multivariate Specifica-tion (p-value) | Bivariate Specifica-tions (p-value) | Multivariate Specifica-tion (p-value) | Bivariate Specifica-tions (p-value) | Multivariate Specifica-tion (p-value) |
| Percent of students scoring "not proficient" in either math or reading/ELA at baseline | 0.06 (0.54) | 0.11 (0.44) | -0.01 (0.88) | 0.01 (0.88) | -0.00* (0.03) | 0.00 (0.15) | 0.06 (0.71) | 0.08 (0.66) | 0.09 (0.71) | 0.27 (0.32) | 0.29 (0.15) | 0.28 (0.29) | 0.01 (0.98) | 0.02 (0.96) | 0.16 (0.10) | 0.17 (0.19) | -0.08 (0.49) | -0.10 (0.49) |
| ***Percent of Control Group Students Receiving Mentoring*** | 0.04 (0.84) | 0.12 (0.62) | -0.11 (0.41) | 0.03 (0.86) | 0.01 (0.77) | 0.01 (0.75) | -0.17 (0.63) | 0.01 (0.97) | 0.42 (0.40) | -0.11 (0.82) | 0.07 (0.86) | 0.14 (0.77) | -0.38 (0.48) | 0.13 (0.83) | 0.06 (0.75) | 0.09 (0.69) | 0.09 (0.69) | 0.04 (0.87) |

\*    p-value<.05, two-tailed test

**Exhibit G.20**

**Site-Level Associations: Relationship between Program Characteristics and Delinquent Behaviors/Participation in Harmful Activities**

| | Misconduct (Student Survey) | | Delinquency (Student Survey) | | Truancy Rate | | Any Misconduct (School Records) | | Repeated Misconduct (School Records) | | Any Delinquency (School Records) | | Repeated Delinquency (School Records) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Bivariate Specifications (p-value) | Multivariate Specification (p-value) | Bivariate Specifications (p-value) | Multivariate Specification (p-value) | Bivariate Specifications (p-value) | Multivariate Specification (p-value) | Bivariate Specifications (p-value) | Multivariate Specification (p-value) | Bivariate Specifications (p-value) | Multivariate Specification (p-value) | Bivariate Specifications (p-value) | Multivariate Specification (p-value) | Bivariate Specifications (p-value) | Multivariate Specification (p-value) |
| *Program Delivery (based on pre-intervention activities or characteristics of mentors)* | | | | | | | | | | | | | | |
| Average hours of mentor pre-match training provided to mentors | 0.01 (0.50) | 0.01 (0.16) | 0.00 (0.50) | 0.00 (0.82) | 0.06 (0.39) | 0.07 (0.35) | 0.01 (0.41) | 0.00 (0.76) | 0.00 (0.64) | -0.01 (0.31) | 0.01 (0.21) | 0.01 (0.25) | 0.00 (0.46) | 0.01 (0.30) |
| Percent of mentors 22 years or below | 0.02 (0.75) | 0.02 (0.75) | 0.02 (0.52) | -0.01 (0.81) | 0.42 (0.46) | 0.27 (0.66) | -0.04 (0.49) | -0.01 (0.88) | -0.02 (0.73) | -0.03 (0.60) | 0.01 (0.81) | 0.09 (0.19) | 0.01 (0.88) | 0.03 (0.51) |
| Percent of mentor/student matches of the same race/ethnicity | 0.09 (0.27) | 0.04 (0.69) | 0.00 (0.92) | -0.04 (0.44) | -0.60 (0.42) | -0.87 (0.28) | 0.06 (0.44) | 0.00 (0.99) | 0.05 (0.42) | -0.03 (0.72) | 0.01 (0.94) | -0.14 (0.20) | -0.05 (0.25) | -0.03 (0.65) |
| *Program Delivery (based on aggregated mentor reports post-intervention)* | | | | | | | | | | | | | | |
| Amount of ongoing mentor support (average frequency of mentor-supervisor meetings) | -0.04 (0.10) | -0.06 (0.06) | -0.01 (0.48) | -0.02 (0.21) | 0.05 (0.83) | 0.07 (0.78) | 0.03 (0.23) | 0.04 (0.37) | 0.02 (0.33) | 0.03 (0.17) | 0.04 (0.05) | 0.07* (0.03) | -0.01 (0.61) | 0.01 (0.69) |
| Frequency of working on delinquency risk avoidance in student-mentor meetings | 0.04 (0.73) | 0.12 (0.42) | 0.08 (0.22) | 0.12 (0.13) | 0.23 (0.84) | 0.30 (0.79) | 0.01 (0.93) | 0.01 (0.94) | -0.07 (0.44) | -0.03 (0.76) | 0.02 (0.89) | -0.15 (0.31) | -0.13 (0.08) | -0.10 (0.34) |
| Percent of mentor/student matches lasting 6 months or longer | 0.01 (0.94) | 0.04 (0.68) | 0.01 (0.84) | -0.01 (0.83) | -0.73 (0.29) | -0.59 (0.43) | -0.06 (0.53) | -0.12 (0.44) | -0.10 (0.09) | -0.12 (0.12) | 0.09 (0.25) | 0.02 (0.86) | -0.06 (0.18) | -0.04 (0.63) |
| Average total hours of mentor/student meetings per month | 0.01 (0.14) | 0.01 (0.14) | 0.01 (0.06) | 0.01 (0.05) | -0.05 (0.31) | -0.01 (0.83) | 0.01 (0.42) | 0.00 (0.84) | 0.01* (0.02) | 0.00 (0.41) | 0.00 (0.50) | 0.00 (0.68) | 0.00 (0.61) | 0.00 (0.80) |
| *Student Characteristics* | | | | | | | | | | | | | | |
| Percent of students with self-reported delinquent behaviors at baseline | 0.21 (0.24) | 0.07 (0.73) | 0.04 (0.67) | 0.00 (0.96) | -4.16* (0.00) | -3.37* (0.02) | 0.21 (0.34) | 0.15 (0.58) | 0.43* (0.00) | 0.35* (0.02) | -0.11 (0.56) | -0.07 (0.70) | -0.03 (0.79) | -0.05 (0.73) |

**Exhibit G.20**

**Site-Level Associations: Relationship between Program Characteristics and Delinquent Behaviors/Participation in Harmful Activities**

| | Misconduct (Student Survey) | | Delinquency (Student Survey) | | Truancy Rate | | Any Misconduct (School Records) | | Repeated Misconduct (School Records) | | Any Delinquency (School Records) | | Repeated Delinquency (School Records) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Bivariate Specifications (p-value) | Multivariate Specification (p-value) | Bivariate Specifications (p-value) | Multivariate Specification (p-value) | Bivariate Specifications (p-value) | Multivariate Specification (p-value) | Bivariate Specifications (p-value) | Multivariate Specification (p-value) | Bivariate Specifications (p-value) | Multivariate Specification (p-value) | Bivariate Specifications (p-value) | Multivariate Specification (p-value) | Bivariate Specifications (p-value) | Multivariate Specification (p-value) |
| Percent of students scoring "not proficient" in either math or reading/ELA at baseline | -0.02 (0.85) | 0.06 (0.61) | -0.06 (0.17) | -0.08 (0.20) | -0.02 (0.98) | -0.56 (0.47) | 0.02 (0.86) | -0.06 (0.76) | 0.03 (0.63) | -0.11 (0.27) | -0.11 (0.20) | -0.17 (0.23) | 0.06 (0.25) | 0.02 (0.81) |
| ***Percent of Control Group Students Receiving Mentoring*** | 0.01 (0.97) | 0.02 (0.93) | -0.08 (0.36) | -0.05 (0.66) | 3.33* (0.02) | 2.72 (0.10) | 0.16 (0.33) | 0.07 (0.78) | 0.16 (0.17) | 0.15 (0.25) | 0.04 (0.80) | 0.11 (0.53) | 0.01 (0.87) | -0.02 (0.90) |

\* p-value<.05, two-tailed test