The Role of Risk

MENTORING EXPERIENCES AND OUTCOMES FOR YOUTH WITH VARYING RISK PROFILES

Carla Herrera
David L. DuBois
Jean Baldwin Grossman

The Role of Risk

MENTORING EXPERIENCES AND OUTCOMES FOR YOUTH WITH VARYING RISK PROFILES

Carla Herrera
David L. DuBois
Jean Baldwin Grossman

Foreword

This evaluation of the Mentoring At-Risk Youth project was initiated by Public/Private Ventures (P/PV) in October of 2007, adding to an extensive body of research that P/PV had conducted on mentoring over several decades. When P/PV ceased operations in July of 2012, this report was still in development. Our colleagues at MDRC generously agreed to publish the report. The findings and conclusions are solely those of the authors.

About the Authors

Carla Herrera, Ph.D., is an independent consultant who was most recently a senior research fellow at P/PV. Dr. Herrera has extensive expertise in mentoring. She has published numerous reports and articles on school-based, community-based and group mentoring over the past 14 years and led P/PV's impact study of the Big Brothers Big Sisters (BBBS) School-Based Mentoring program. These studies have helped inform the field about the relationships that develop in these programs, the experiences of youth and mentors, how youth benefit and how program practices may shape these experiences and benefits. Her current work includes consulting on a national evaluation led by the American Institutes for Research that examines the effects of various practice enhancements on match success. Dr. Herrera is a member of both the Big Brothers Big Sisters of America (BBBSA) Research Advisory Council and MENTOR's Research and Policy Council. She has a B.A. from Stanford University and a Ph.D. in developmental psychology from the University of Michigan.

David L. DuBois, Ph.D., is a professor in Community Health Sciences within the School of Public Health at the University of Illinois at Chicago. His research examines the contribution of protective factors, particularly self-esteem and mentoring relationships, to resilience and holistic positive development, with a focus on translating this knowledge into the design of effective youth programs. He has led two widely cited meta-analyses of evaluations of the effectiveness of youth mentoring programs. Dr. DuBois is also lead coeditor of the first and second editions of the *Handbook of Youth Mentoring* (Sage Publications 2005 and forthcoming) and coauthor of *After-School Centers and Youth Development: Case Studies of Success and Failure* (Cambridge University Press 2012). He is a fellow of the American Psychological Association and Society for Community Research and Action and a past distinguished fellow of the William T. Grant Foundation and consults widely to mentoring programs nationally and internationally. He received his Ph.D. in clinical-community psychology from the University of Illinois at Urbana-Champaign.

Jean Baldwin Grossman, Ph.D., is a senior research fellow at MDRC, and on the faculty of Princeton University's Woodrow Wilson School. She is an expert on mentoring programs, after-school programs and evaluation design. She has decades of experience developing and conducting evaluations, including 11 random assignment evaluations. While considered an evaluation design expert, her substantive specialty is the study of programs for disadvantaged adolescents, especially mentoring and out-of-school-time programs. She has studied mentoring programs for almost two decades, being intimately involved with the four interrelated studies that comprised P/PV's multi-year multi-site evaluation of the BBBS community-based mentoring program and later the multi-year study of the BBBS school-based program. Prior to working at MDRC, she worked at P/PV and Mathematica Policy Research. She has a Ph.D. in economics from M.I.T.

Other Contributors

Washington State Mentors (WSM) is a public/private partnership that has been serving the state's youth mentoring community since 2004. WSM promotes and supports high-quality mentoring to foster positive youth development and academic success. WSM uses data from its annual statewide mentoring survey to inform the work of mentoring programs, state and local government leaders, and funders. It also conducts a statewide conference and offers a suite of training and technical assistance to programs

throughout the state. WSM served as the intermediary for the Mentoring At-Risk Youth project, under the leadership of **Janet Heubach, Ph.D.** As WSM's senior program officer, Dr. Heubach also leads the statewide Quality Mentoring Assessment Path initiative and the annual State of Mentoring Survey, and develops new evaluation and demonstration projects. She previously worked for the Pacific Northwest National Laboratory as a senior research scientist. Dr. Heubach received her B.A. from the University of Colorado, M.A. from the University of Wyoming and Ph.D. from the University of Washington.

Chelsea Farley, M.P.H., specializes in developing clear and effective communications for nonprofits, foundations and research organizations. She has written and edited numerous influential papers and reports, focused mostly on education and youth programs, workforce development and the criminal justice system. Ms. Farley earned her B.A. in sociology from Wesleyan University and her M.P.H. from Boston University.

Michael J. Karcher, Ed.D., Ph.D., is a professor of counseling at the University of Texas at San Antonio, where he coordinates the School Counselor Training Program. He conducts research on school-based and cross-age peer mentoring as well as on adolescent connectedness and pair counseling. He is on the editorial board for several national journals and the research and advisory boards of BBBSA and MENTOR. He received doctorates in human development and psychology from Harvard University and in counseling psychology from the University of Texas at Austin.

Daniel A. Sass, Ph.D., is assistant professor of educational psychology and director of the Statistical Consulting Center in the Department of Management Science and Statistics at the University of Texas at San Antonio. His research interests include methodological issues related to multivariate statistics, with a central focus on factor analysis, structural equation modeling and instrument development. Dr. Sass earned his Ph.D. in educational psychology with an emphasis in educational statistics and measurement from the University of Wisconsin-Milwaukee.

Acknowledgments

Contents

Executive Summary

Introduction

More and more, mentoring programs are being asked to serve higher-risk youth—for example, those in foster care or the juvenile justice system or youth with a parent who is incarcerated.[1,2] This impulse is understandable: Studies have illuminated the varied benefits that mentoring programs can provide, including improving academics and relationships with others and reducing involvement in problem behaviors.[3] Higher-risk youth are clearly in need of such support.

While these youth are often viewed through the lens of likely future costs to their communities, they also embody enormous unrealized potential. With the right kinds of support, these young people could put themselves on a path toward bright, productive futures, and make vital contributions to their families, neighborhoods and nation. Many hope that mentoring programs can help make this vision a reality. Yet few studies have examined and compared the benefits of mentoring for youth with differing types or sources of risk.

The Role of Risk: Mentoring Experiences and Outcomes for Youth with Varying Risk Profiles presents findings from the first large-scale study to examine how the levels and types of risk youth face may influence their relationships with program-assigned mentors and the benefits they derive from these relationships. The study looked closely at the backgrounds of participating youth and their mentors, the mentoring relationships that formed, the program supports that were offered, and the benefits that youth accrued—and assessed how these varied for youth with differing "profiles" of risk. We believe the study's results provide useful guidance for practitioners, funders and policymakers who want to know which youth are best suited for mentoring and how practices might be strengthened to help ensure that youth facing a variety of risks get the most out of their mentoring experience.

This summary highlights the major findings and implications from the full report, which is available at www.mdrc.org and www.wamentors.org.

Key Findings from the Study

This study examined mentoring program relationships, experiences and benefits for higher-risk youth, with five key findings:

- Without substantial effort beyond their normal outreach strategies, programs were able to reach and serve youth facing a wide range of challenges.

- Youth with differing risk "profiles" (that is, levels and types of risk) had relationships of similar strength and duration and derived similar benefits from program participation.

- However, the challenges reported by mentors and the reasons matches ended differed as a function of youth's risk profile.

- The strongest program benefit, and most consistent across risk groups, was a reduction in depressive symptoms—a particularly noteworthy finding given that almost one in four youth reported worrisome levels of these symptoms at baseline. Findings also suggested gains in social acceptance, academic attitudes and grades. Youth did not appear to benefit in their relationships with parents or in their positive or negative behaviors.

- Mentors who received early-match training and consistent program support met more frequently and had longer-lasting relationships with their mentees. Youth whose mentors received training also reported higher-quality relationships.

About the Study

In 2007, the Bill & Melinda Gates Foundation commissioned an independent evaluation to examine the services and effectiveness of mentoring programs for youth with different profiles of risk. Washington State Mentors (WSM) served as the project's intermediary, providing implementation oversight and support to participating programs. WSM selected seven mentoring programs serving youth in Washington State[4] to participate in the initiative. All the programs utilized volunteers to provide one-to-one mentoring to youth in community settings.[5] Five of them were operated by Big Brothers Big Sisters agencies.

The programs were asked to reach out to "higher-risk" youth—that is, youth who faced significant personal and/or environmental challenges. We then collected information about:

- *Youth risk*, through a detailed survey administered to parents at enrollment;

- *Youth outcomes*, through surveys administered to youth (and, for a portion of the sample, parents) first at program enrollment and then again at a 13-month follow-up;

- *Mentor program experiences*, through surveys completed by mentors;

- *Mentoring relationship quality*, through surveys of mentors and youth;

- *Mentor-youth meetings, match duration and program supports*, through program records and surveys completed by the supervisors of case managers; and

- *Program practices*, through surveys administered to program staff.

Assessing Youth Risk

The study's approach to assessing risk drew on past research suggesting that both "environmental" risk (that is, challenges in the youth's surrounding life circumstances, such as poverty or living in a dangerous neighborhood) and "individual" risk (that is, challenges in the young person's behavior, social or academic functioning, or health) may shape the extent to which youth benefit from mentoring.[6] Based on a survey of parents, we categorized youth into four distinct risk profiles:

- Youth relatively high on both individual and environmental risk (the "highest-risk" youth in the sample),

- Youth with relatively low individual but high environmental risk,

- Youth with relatively high individual but low environmental risk, and

- Youth relatively low on both types of risk (the "lowest-risk" youth in the sample).

Who Did the Programs Reach?

All participating programs were asked to reach out to "higher-risk" youth. One of the issues we set out to explore was how successful they were in these efforts.

The seven programs reached youth facing a wide range of challenges—without significant effort beyond their normal outreach strategies. The programs enrolled 1,310 youth in the study, who ranged in age from 8 to 15 (and averaged a little over 11 years old). About half were male, and 57 percent were ethnic minorities. As in many mentoring programs around the country, a large proportion of the youth came from single-parent homes (about two thirds) and low-income households (about two fifths had annual incomes below $20,000). In addition, nearly three quarters (71 percent) faced some type of "individual-level" risk—for example, academic struggles, behavior problems or mental health concerns.

Considering various criteria and definitions, the youth in the study are, as a whole, best categorized as "higher risk" rather than "high risk." As a group, the youth in our study were much more likely to

face a number of risk factors than the average child in the US, but few had engaged in behaviors like substance use or crime that are often used to determine "high-risk" status, perhaps in part because they were fairly young. Thus, overall, the youth in the study are best thought of as "higher risk"—a designation that falls somewhere between what would typically be characterized as "at risk" and "high risk." However, there was substantial variability in both the levels and types of risk that these youth experienced.

What Kinds of Relationships Did Youth Experience?

The relationships between youth and their mentors are the central route through which mentoring is generally thought to benefit young people,[7] and research has linked stronger and longer mentoring relationships to more favorable youth outcomes.[8] In this study, we found that:

Mentors and youth reported fairly strong relationships. We explored three aspects of youth-reported relationship quality: 1) closeness; 2) the extent

to which the relationship included opportunities for learning and working toward goals; and 3) the extent to which the mentor considered the youth's interests and input. Almost three quarters of youth reported at least a moderately positive relationship with their mentor across all three of these dimensions. Mentors, on average, also reported fairly strong feelings of closeness toward their mentee.

Yet, almost half of the youth had experienced at least one match closure by the time of our 13-month follow-up survey. Some of these youth had been rematched, yielding an average of 9.6 total months of mentoring across all matches.[9] Still, overall, only about 60 percent of participants were in an active match at follow-up. Mentors reported initiating the end of the match more than half of the time. Two of the most common reasons cited by mentors were "not enough youth interest" (33 percent) and, similarly, the impression that the youth did not seem to need a mentor (17 percent). Despite the serious challenges faced by many of these youth, in only about 10 percent of cases did mentors report that the match closed because the youth's needs were too severe.

Importantly, match quality and length did *not* vary notably based on the youth's risk profile. The frequency of meetings between youth and their mentor and the total number of hours the match met throughout the study period were also, for the most part, consistent across the risk groupings.

The similarities in relationship quality and duration across the risk groups belie very different challenges and reasons why matches ultimately ended. For example, mentors who were matched with youth who were relatively high on individual risk were more likely to report significant challenges with their mentee's behavior. In contrast, mentors matched with youth high on environmental risk were more apt to report challenges connecting with and getting support from the mentee's family as well as frequent cancellations of match meetings by youth. Mentors matched with the lowest-risk youth were most likely to report relationships ending due to a lack of youth interest or the youth not seeming to need a mentor.

Assessing Youth Outcomes

The evaluation's design allowed us to assess the effects of mentoring program participation in two ways:

- **Experimental/Random Assignment Component:** In the first year of the evaluation, in the two largest programs, about half of the youth were randomly selected to be matched immediately with mentors (the "treatment group"), while the remaining half (the "control group") were not eligible for matching until after the study's 13-month follow-up assessment. To assess impacts, we compared the change over time in the outcomes of youth in the treatment group to that in the control group.

- **Quasi-Experimental Component:** In the other five programs and during the second year at the two largest programs, all eligible youth were enrolled in the evaluation and offered a mentor. In this study component, we compared the change over time in the outcomes of all youth who were offered a mentor without going through random assignment to that in the control group from the random assignment portion of the study (in this context, referred to as a "comparison group").

How Did Youth Benefit?

Findings suggest that mentoring benefited youth's emotional/psychological well-being, peer relationships, academic attitudes, and grades. At the 13-month follow-up assessment, findings from the quasi-experimental portion of the evaluation indicated that mentored youth were doing significantly better than youth in the non-mentored comparison group on a number of important outcome measures. In particular, these youth reported:

- Fewer depressive symptoms;

- Greater acceptance by their peers;

- More positive beliefs about their ability to succeed in school; and

- Better grades in school.

We also wanted to assess whether mentored youth did better *overall* across the set of outcomes we tested. Mentoring is believed to address the distinct needs of participating youth, suggesting that only

some youth may benefit in any particular area (a gain that might be missed when examining change in individual outcomes across an entire group). Thus, we developed a measure of aggregate positive change for this study and found that mentored youth in fact showed meaningful improvement in a greater number of our key outcomes than youth in the comparison group.

In the random assignment portion of the study, we found evidence of significant benefits for only two outcomes: depressive symptoms and the aggregate measure of positive change. Because these two impacts were found in both components of the evaluation, we believe the study provides particularly strong evidence about the programs' benefits in these areas. The evidence for mentoring's ability to influence academics and peer relationships is more moderate.

Program benefits were not evident in either portion of the evaluation for the other outcome measures we assessed:

- Positive behavior toward peers;

- Skipping school;

- Misconduct; or

- Parent trust.

Youth also did not differ on our aggregate measure of the number of outcomes for which there was evidence of negative change.

Did Impacts Vary by Youth's Risk Profile or Other Background Characteristics?

Overall, program benefits were fairly similar for youth regardless of their risk profile and other background characteristics. Indeed, youth in all four risk groups appeared to derive at least some gains from their participation. The study's findings as a whole thus suggest that the benefits of volunteer-centered community-based mentoring are not confined to youth with particular types or levels of risk. There were some exceptions to this general pattern—most notably a trend toward somewhat stronger and more consistent benefits for youth who were relatively high on individual but not environmental risk.

How Were the Matches Supported?

Programs varied in the types and amount of support they offered to participating matches. And even within each program, matches varied in their experience of key supports—for example, how much training mentors received and the extent to which they felt training and support were sufficient. As part of their involvement in the study, three programs also implemented specific enhancements that were designed to increase the support available to matches. When we examined various program practices, we found that:

Matches received fairly similar types and levels of support regardless of youth's risk status, with one notable exception. Mentors paired with youth who were relatively high on individual risk were more likely to have had early-match training and regular support contacts with program staff. They also reported lengthier support calls.

Mentors' self-reported training/support *needs* did differ markedly depending on their mentee's risk profile. For example, mentors paired with the highest-risk youth were more likely to say they needed help learning how to interact with the youth's family or navigating social service systems, while those whose mentees were high on individual risk reported greater concerns about dealing with youth's social and emotional issues.

The supports received by mentors, parents and youth were linked with key match outcomes. Mentors who received early-match *training*[10] met more frequently with their mentee and were more likely to have a match that lasted at least 12 months. In addition, youth paired with these mentors rated their mentoring relationship as being of higher quality. *Regular support calls from case managers to mentors* were also linked with longer-lasting matches and more frequent meetings between mentors and youth. The findings suggest that the *quality* of case manager support was important as well, contributing to both the strength and longevity of the match. Finally, matches in which parents and youth received regular support calls from case managers met more frequently than matches without this level of support.

Implications for Practitioners and Funders

The findings from this study have several noteworthy implications for practitioners and funders:

1. **Training and support for matches should be tailored to the types and levels of risk experienced by youth.** We found significant differences in the challenges and support needs that mentors recounted, based on their mentee's risk profile. Although matches involving higher-risk youth seemed to present greater challenges, all matches, including those with the lowest levels of risk, brought distinct issues and concerns. This highlights the need to tailor program training and support to the specific levels and types of risk faced by participating youth. To do this effectively, programs will need to systematically assess youth risk at intake, gathering information about difficulties in the youth's environment and about personal challenges, such as behavior problems or mental health issues. Funders should support programs' efforts to better measure youth risk and to tailor the training and support they offer accordingly.

2. **Mentoring should be broadly available, as youth with varying levels and types of risk appear to derive important benefits.** Overall, the study did not find strong evidence that mentoring benefited youth differently based on their risk profile or other background characteristics. These findings argue against restricting eligibility or recruitment efforts to youth with particular risk profiles or backgrounds, at least for programs that are structured similarly to the ones in this study. At the same time, for programs interested in targeting higher-risk youth, the study's findings provide optimism that such youth can be recruited and that, with the right supports in place, these youth can derive significant benefits from mentoring.

3. **Greater emphasis should be placed on the mental health needs of youth and the benefits that mentoring can provide in this area.** Depression has been linked to a host of short- and long-term problems for young people, including suicidal behavior, academic and social difficulties, and increased risk for substance abuse and teen pregnancy.[11] It was striking that almost one in four youth in this study reported high levels of depressive symptoms at baseline. Our findings offer robust evidence that participation in mentoring programs can ameliorate and/or prevent the emergence of depressive symptoms. This is highly encouraging, given the number of other areas (personal, social and academic) that may benefit from better mental health. One key implication for programs is the importance of careful screening for mental health issues, both at intake and over the course of a young person's involvement in the program, in combination with referral mechanisms for youth who are in need of additional support. At the funding level, the findings from this study suggest that mental health outcomes should be given greater weight in designing and evaluating the success of mentoring initiatives.

4. **Efforts should continue to improve the strength and consistency of the benefits that youth derive from mentoring programs.** As a whole, the findings of this study point to a positive, but somewhat inconsistent pattern of benefits for youth who had access to volunteer-centered, one-to-one community-based mentoring over a 13-month period. For example, the evaluation found no evidence that mentoring helped to curb youth involvement in problem behavior. This aspect of the study's results underscores a need for moderation when forecasting the likely impact of mentoring as an intervention strategy.[12] The findings also suggest, however, that by improving program supports (such as the training provided to mentors or to the staff who support the matches), it may be possible to strengthen mentoring relationships and potentially, in turn, increase the impact of program involvement on youth outcomes. Funding support will be necessary to make large-scale in-roads in this area. These efforts should include support for intermediary organizations that can broker needed technical assistance and bring programs together to share lessons about effective practice.

While these caveats are important to keep in mind, we believe the findings from the study support an optimistic outlook about the role that mentoring programs can play in the lives of youth facing a wide variety of risks—including those who are often deemed "hardest to serve" in social programs (that is, those who are relatively high on *both* environmental and individual risk). In sum, the high hopes that policymakers and funders have had for mentoring programs serving higher-risk youth may

be well founded, particularly if programs continue to refine their efforts to ensure that matches get the targeted training and support they need.

Endnotes

1. For an overview of recent developments in the youth mentoring field that illustrate the growing focus on serving higher-risk youth, see DuBois, D. L. and M. J. Karcher. In press. "Youth Mentoring in Contemporary Perspective." To appear in D. L. DuBois and M. J. Karcher (eds.). *Handbook of Youth Mentoring.* (2nd ed.). Thousand Oaks, CA: Sage.

2. In this report, "higher-risk" youth also include youth with other types of serious challenges, such as mental health difficulties and those experiencing relatively large numbers of risk factors across different life domains.

3. For example, see Tierney, J. P. and J. B. Grossman with N. L. Resch. 1995. *Making a Difference: An Impact Study of Big Brothers Big Sisters.* Philadelphia: Public/Private Ventures. See also DuBois, D. L., B. E. Holloway, J. C. Valentine and H. Cooper. 2002. "Effectiveness of Mentoring Programs for Youth: A Meta-Analytic Review." *American Journal of Community Psychology*, 30 (2), 157–197. Also, DuBois, D. L., N. Portillo, J. E. Rhodes, N. Silverthorn and J. C. Valentine. 2011. "How Effective Are Mentoring Programs for Youth? A Systematic Assessment of the Evidence." *Psychological Science in the Public Interest*, 12 (2), 57–91.

4. One program was based in Oregon, but agreed that at least 80 percent of participating youth would live in adjacent Washington communities that also were part of its service area.

5. In one program, mentors did not meet with youth on their own in the community. However, because the matches met in a variety of settings (at the university that houses the program, in youth's schools, and at monthly family activities), it was characterized as, at least in part, a community-based mentoring program.

6. DuBois et al. 2002, 2011. Op cit.

7. Rhodes, J. E. 2005. "A Model of Youth Mentoring." In D. L. DuBois and M. J. Karcher (eds.). *Handbook of Youth Mentoring*. Thousand Oaks, CA: Sage, 30–43.

8. Grossman, J. B. and J. E. Rhodes. 2002. "The Test of Time: Predictors and Effects of Duration in Youth Mentoring Programs." *American Journal of Community Psychology*, 30, 199–219. Herrera, C., J. B. Grossman, T. J. Kauh, A. F. Feldman and J. McMaken, with L. Z. Jucovy. 2007. *Making a Difference in Schools: The Big Brothers Big Sisters School-Based Mentoring Impact Study*. Philadelphia: Public/Private Ventures. Rhodes, J. E. and D. L. DuBois. 2006. "Understanding and Facilitating the Youth Mentoring Movement." *Social Policy Report: Giving Child and Youth Development Knowledge Away*. Available from Society for Research in Child Development at www.srcd.org/sites/default/files/documents/20-3_youth_mentoring.pdf.

9. This total reflects the amount of time the youth had been mentored at the time of their follow-up survey. Many of these matches were ongoing, however. Thus, they ultimately lasted longer than what we measured at follow-up.

10. "Training" in the initiative needed to have several characteristics, including being interactive and curriculum-based; thus, a basic orientation to program guidelines and expectations would not qualify as training.

11. See Malhotra, S. and P. P. Das. 2007. "Understanding Childhood Depression." *Indian Journal of Medical Research*, 125, 115–128. See also Cash, S. J. and J. A. Bridge. 2009. "Epidemiology of Youth Suicide and Suicidal Behavior." *Current Opinion in Pediatrics*, 21 (5), 613-619.

12. Wheeler, M., T. Keller and D. L. DuBois. 2010. "Review of Three Recent Randomized Trials of School-Based Mentoring." *Social Policy Report*, 24, 1–21. Available from Society for Research in Child Development at www.srcd.org. DuBois et al. 2011. Op cit.

Introduction

Chapter I

Young people need a range of caring adults in their lives to be successful. Yet, as many as 1 in 5 youth—and even more of those living in poverty—lack this vital resource (America's Promise Alliance 2006). Mentoring programs represent one promising avenue for helping to meet this need.

In 1995, Public/Private Ventures (P/PV) published its landmark random assignment impact study of the Big Brothers Big Sisters (BBBS) community-based mentoring program. That study provided rigorous evidence that formal, well-run volunteer-centered, community-based one-to-one mentoring programs could yield numerous benefits for youth—in academic, social and behavioral domains (Tierney et al. 1995). More recent studies generally have supported these findings, providing further evidence that mentoring programs can improve key developmental outcomes for youth across a wide variety of areas.[1]

The youth included in most evaluations have been those typically targeted by mentoring programs—youth whose experiences or circumstances (for example, living in a high-crime neighborhood or having difficulties succeeding in school) put them "at risk" for future problems. Some of these youth have very specific needs and face what typically would be seen as more serious risks (such as child abuse, homelessness or involvement in the juvenile justice system). However, programs by and large have not served such youth in large numbers.

In recent years, there has been growing concern about "higher-risk" youth—those who are particularly vulnerable to problems as they get older, including school dropout, teen parenthood, mental health issues and crime. Indeed, these higher-risk youth are often viewed through the lens of likely future costs to their communities, but they also embody enormous unrealized potential. With the right kinds of support, these young people could put themselves on a path toward bright, productive futures, and make vital contributions to their families, neighborhoods and nation.

Describing Youth "Risk"

In this report, we use the term "at risk" to describe youth with one or more characteristics, behaviors or features of their surrounding environments that research has associated with an increased likelihood of future problems, such as academic failure, unemployment or crime. We use the term "higher risk" to describe those youth with more numerous challenges in their lives, but who are not necessarily at the extreme end of the risk continuum. These youth, without intervention, have a greater likelihood of negative outcomes than at-risk youth. For more detail about these terms, see Chapter 2.

It makes sense then that policymakers, funders and community leaders want to identify these youth early and get them on a positive track *before* serious problems emerge. As a result, more and more programs are being asked to target some of their services to specific higher-risk groups of youth—for example, youth in foster care or the juvenile justice system or youth with a close family member who is incarcerated (DuBois, Karcher in press). Mentoring is a particularly attractive intervention for these youth given its low cost relative to more intensive programs, the wide variety of areas it appears to benefit (DuBois et al. 2002, 2011) and its flexibility in serving youth with a wide range of backgrounds (see DuBois et al. 2011).

As it stands, however, we know very little about mentoring programs' capacities to serve higher-risk youth. Have programs already started reaching these youth in significant numbers? Is finding and serving these youth challenging? Can mentoring programs actually yield benefits for higher-risk youth that are comparable to the benefits seen in "typical" at-risk mentees? And is there a level of risk at which mentoring is most helpful—or past which benefits fail to accrue? What kinds of adaptations might be needed for programs to serve higher-risk youth effectively?

Intuitively, it would seem that youth *without* significant risk factors may be more receptive to having a mentor than their higher-risk peers. For example, they may be more accustomed and open

to supportive relationships with adults and have stronger interpersonal skills to help the relationship flourish. Mentors also may be better prepared to meet the needs of these youth. At the same time, because higher-risk youth face more challenges personally (like poor school performance) and/or in their day-to-day lives (for example, stressed family relationships), they may have a greater need for a mentor and more room for gains as a result of their mentoring experience.

Few studies have examined and compared benefits of mentoring program participation across groups of youth with differing types or sources of risk. In preliminary attempts to explore this issue, DuBois and his colleagues used a technique known as meta-analysis—which involves analyzing results from multiple past studies—to examine how the magnitude of mentoring programs' effects may have been related to the levels and types of risk faced by participating youth (DuBois et al. 2002, 2011). Interestingly, their findings suggest that higher-risk youth have benefited from program involvement at least as much as lower-risk youth. The differences that did emerge, furthermore, were dependent on the *type* of risk to which youth had been exposed—namely, whether youth had experienced risks at the "individual" level (that is, challenges in the youth's behavior, social or academic functioning or health) or at the "environmental" level (that is, challenges in the youth's surrounding environment, such as living in poverty or a single-parent home), or some combination of the two.

The current study builds on this earlier research—which was limited to examining program effects using the *average* "risk profile" across all participants in a program. By contrast, the current study investigates the risk backgrounds of *individual* youth and how their experiences in the program, as well as its impact, may differ based on the levels and types of risk they face. No other large-scale mentoring evaluation has collected and systematically analyzed as much information about risk, asking families in-depth questions about their children's home life, peers, behavior, academics, mental health and more.

The study also looked closely at the services provided by programs (for example, were they similar or different across youth with varying profiles of risk, and how could they be strengthened overall?), at the mentoring relationships themselves (for instance, how did the strength or length of the mentoring matches differ depending on youth's risk status?) and at mentor experiences and challenges (for example, in what areas did mentors need program support, and did these vary based on the risk profiles of youth?). By examining these issues, the study provides today's mentoring programs—especially those that utilize volunteers and a one-to-one approach—with the strongest research yet to determine which groups of youth they are best suited to serve and how they could strengthen their mentoring models to ensure they are providing *all* participating young people with the best possible support.

Requirements for Programs in the Initiative

- Youth met one-to-one with mentors in the community;[a]

- Program served youth between the ages of 9 and 14;

- Program expected matches to meet a minimum of twice a month for four or more hours over at least a 12-month period;[b]

- The number of matches targeted for study participation could not exceed 20 percent of the program's total number of active community-based matches before the study began;

- Program had a minimum annual budget of $100,000;

- Program had at least 100 active one-to-one, community-based matches;

- Program's practices met all standards and benchmarks from the *Elements of Effective Practice for Mentoring* (MENTOR, 2009); and

- Program had been operating for at least two continuous years.

[a] In one program, mentors did not meet with youth on their own in the community. However because the matches met in a variety of settings (at the university that houses the program, in youth's schools and at monthly family activities), it was characterized as, at least in part, a community-based mentoring program.

[b] Programs could have expectations for match meetings and length that went beyond the minimum guidelines.

Study Overview

In 2007, the Bill & Melinda Gates Foundation commissioned an independent evaluation to examine the services and effectiveness of mentoring programs for youth with different profiles of risk. Washington State Mentors (WSM) served as the project's intermediary, providing implementation oversight and support to participating programs.[2] This report presents the findings from the evaluation, including answers to the following questions:

- Can mentoring programs reach higher-risk youth? To what extent are they already serving youth with multiple or more serious risk factors?

- Does the quality and/or length of mentoring relationships differ for youth with varying profiles (that is, levels and types) of risk?

- What kinds of programmatic supports are needed to enable mentoring to succeed with youth of varying risk profiles?

- Is mentoring effective with higher-risk youth? Does mentoring's impact differ by level or type of risk, and if so, how?

WSM selected seven mentoring programs serving youth in Washington State[3] to participate in the initiative.[4] The participating programs ranged widely in the number of youth they served and included one that focused on working with higher-risk youth (see Table 1.1). None were stand-alone "grassroots" organizations; that is, all were affiliated with larger national nonprofits or, in one case, a university. In addition to several broad requirements that all programs had to meet to be part of the initiative (listed in the Requirements for Programs in the Initiative text box on the previous page), the two largest programs were asked to implement several programmatic enhancements over the course of the study. These enhancements (which are described in more detail in Chapter 4) were designed to increase the chances that higher-risk youth would benefit, and they provided the study with a wide range of practices to assess across all seven programs. This variation enabled us to investigate whether certain practices, including these enhancements, were associated with stronger and more enduring mentoring relationships.

Table 1.1
Program Characteristics

At the start of the study…	Program 1	Program 2	Program 3	Program 4	Program 5	Program 6	Program 7
Number of years the program had provided mentoring services	43	4	12	20	7	48	31
Number of youth matched with community-based mentors in the previous year	391	41	100	305	1,650	2,110	176
Big Brothers Big Sisters agency?	Yes	No	No	Yes	Yes	Yes	Yes
Focus on higher-risk youth?	No	Yes	No	No	No	No	No
Study matches met in school in addition to required community-based meetings?	No	No	Yes	No	No	No	No
Implemented the full set of program enhancements?[a]	Yes	No	No	No	Yes	Yes	No

[a] Only Programs 5 and 6 were required to implement the set of program enhancements as part of their involvement in the initiative. And this was the case only for youth who were assigned to the treatment (mentoring) group in the random assignment evaluation that was conducted at these programs during the first year of the study. Program 1 also chose to implement the enhancement practices. Other participating programs were implementing components of the enhancements as part of their regular practices, but did not choose to implement the full set of enhancements.

Study design. All programs were asked to reach out to higher-risk youth (that is, youth who faced personal and/or environmental challenges). In the first year of the study, in the two largest participating programs, about half of the youth were randomly selected to be matched immediately with mentors (the "treatment group"), while the remaining half (the "control group") were not eligible for matching until 13 months into the study.[5] In the other five programs and during the second year at the two largest programs, all youth who were enrolled in the study were offered a mentor. (See Figure 1.1.)

We assess the effects of mentoring in Chapter 5 in two distinct ways. First, we compare the outcomes of all youth randomly assigned to the treatment group with those randomly assigned to the control group. These youth are only from programs 5 and 6 and were all enrolled in the first year of the study. This constitutes the experimental, or random assignment, component of the study's assessment of program benefits. Second, we compare the outcomes of all youth who were assigned a mentor without going through random assignment (that is, the "participant group") with those of youth in the control group at the two programs participating in the random assignment part of the evaluation (in this context, this non-mentored group is referred to as the "comparison group"). The youth in this study component were enrolled in all seven programs and were enrolled in both the first and second year of the study.[6] This constitutes the quasi-experimental component of the study. The changes in outcomes of the non-mentored comparison group were used to estimate how the treatment or participant group would have changed over 13 months (for example, through normal development) if they had not received mentoring. In general, when the treatment or participant group's scores on an outcome improved significantly more (or worsened significantly less) than those of the comparison group, we concluded that mentoring had a favorable effect on that outcome.[7]

Figure 1.1:
Study Design Used to Assess Program Effects on Youth Outcomes
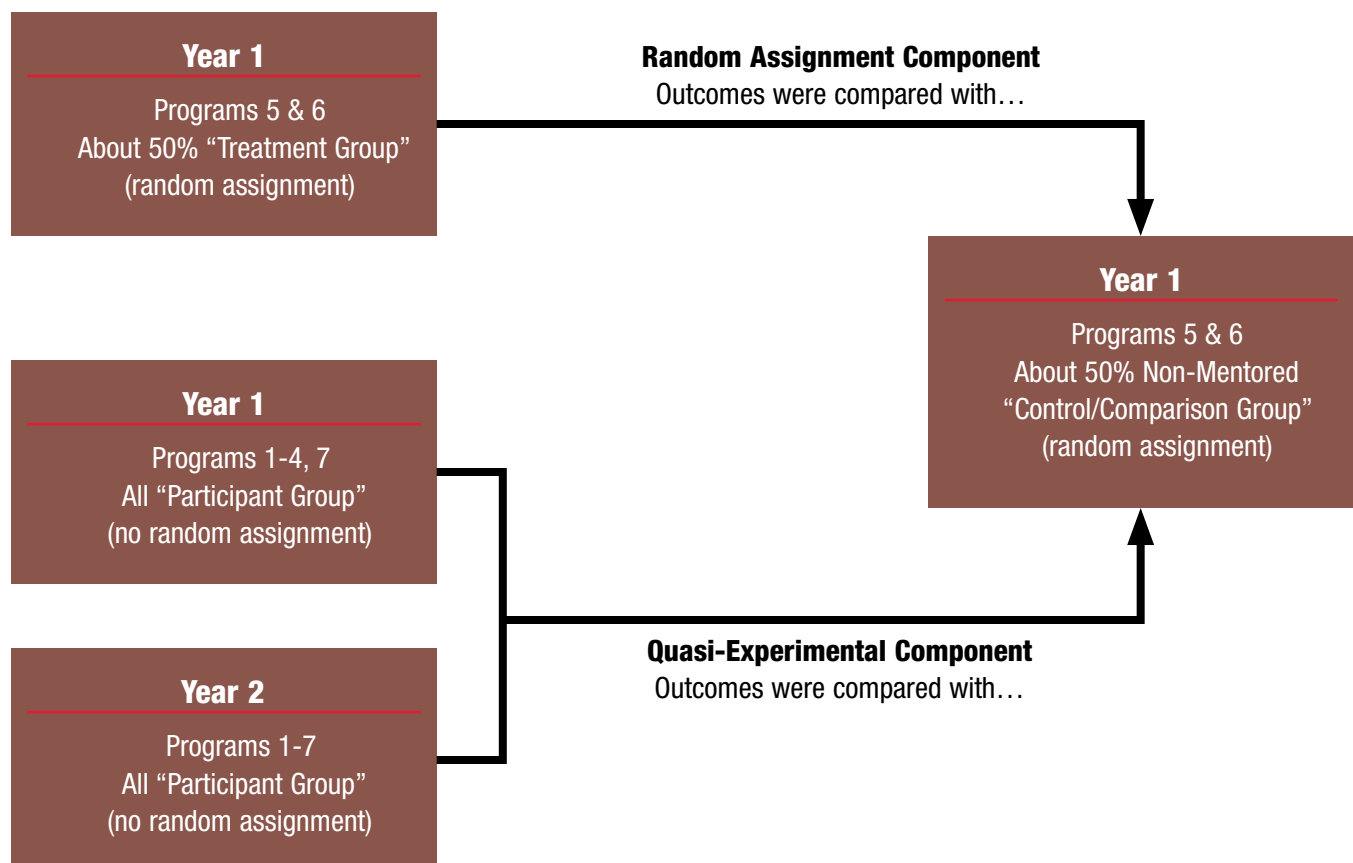
Although the comparison group provides a similar group with which to compare the progress of youth in the participant group (to determine what they would look like without mentoring), the comparison group came from only two of the seven programs and were all enrolled in the first year of the project (a year which may have differed from the following year and affected families' progress in different ways).[8] Thus, the conclusions we can make using the quasi-experimental component of the study are not as definitive as can be made using the experimental component in which youth were randomly assigned to receive mentoring or not. However, by comparing outcomes across both samples, we can assess whether we have strong evidence of program impacts (in cases where both studies yield similar findings) or more moderate evidence (in cases where only one yields significant findings).

In other chapters of the report, which are not concerned with estimating program outcomes, we combine all youth who received a mentor through the initiative. We refer to this combined group as "program participants" (that is, the group of youth who participated in the program) and distinguish it from all those who participated in the study (including those youth in the comparison group), whom we refer to as "study participants."

Data collection. All youth enrolled in the initiative over the two-year study enrollment period were asked to complete a baseline survey at intake as well as a follow-up survey 13 months later. The main outcomes examined in the study were assessed through youth's responses to questions in the surveys they completed at baseline and follow-up (discussed in Chapter 5). As part of the follow-up survey, youth who had been matched with a mentor answered questions about the quality of that relationship (described in Chapter 3).

Each youth's parent was also asked to complete a baseline survey at intake, which elicited information about the risks faced by the youth along with basic demographics for the youth and his or her family (discussed in Chapter 2).

We also collected surveys from mentors at 13 months (or when the match ended, if it ended before 13 months). Mentors were asked questions about their relationship with their mentee, the activities they engaged in, the challenges they faced in their match, whether and why the relationship had ended (discussed in Chapter 3), and their experience of program supports (presented in Chapter 4).

Programs completed surveys about their practices (for example, mentor recruitment, training and support) and characteristics of the mentors who volunteered in their programs and the youth whom they served (for example, gender, ethnicity, family background) at the beginning, middle and end of the study. In addition, programs recorded detailed information on each participating match throughout the study. These data allowed us to track the extent to which programs implemented a range of practices (discussed in Chapter 4), including pre- and post-match mentor training and support calls to parents, youth and mentors as well as how often and for how long matches met each month and match start and end dates. Finally, we asked programs for information about the staff who provided supervision and support for each match. This included information on their education and training as well as the supervisor's ratings of staff effectiveness in supporting matches. (See Appendix A for more details on the study's design and methodology.)

Strengths and Limitations of the Study's Design

Several strengths and limitations of the evaluation should be noted. First, the study includes a random assignment component, which lends strength to our analyses of program outcomes. However, our conclusions about benefits are more tentative than would be possible with a full-scale random assignment design. Second, we had very little past research to rely on in testing associations between youth risk and the characteristics of their relationships. As such, our risk analyses were meant to explore these issues and set the stage for future research rather than to confirm or disprove specific hypotheses. Using this approach, a large number of statistical tests were conducted to assess differences in program benefits and experiences based on youth's risk profiles. Thus, some of the differences we noted may reflect chance fluctuations that would not be found in another sample. Finally, our sample did not include large numbers of very high-risk youth (for example, youth involved with

the juvenile justice system). The experiences and benefits outlined thus may not generalize to such youth. Despite these limitations, we believe the evaluation provides a crucial first step in exploring in detail how risk may shape mentoring relationships and youth outcomes and in outlining how programs may need to approach working with youth of varying risk profiles.

Structure of the Report

In the next chapter, we describe the youth participating in the initiative and the risk factors they faced. In Chapter 3, we examine characteristics of the mentoring relationships these youth experienced— the activities they engaged in with their mentors, the length and quality of the relationships, the challenges mentors reported—and how these characteristics were associated with the levels and sources of risk that youth were experiencing. In Chapter 4, we describe the practices implemented by programs to support the mentoring relationships. Chapter 5 examines the benefits experienced by youth and if and how they varied depending on youth's risk profiles. In the final chapter, we present our conclusions and recommendations for funders and practitioners.

Who Did the Programs Reach?

Chapter II

There is little doubt that many youth could benefit from a positive, supportive relationship with an adult outside of their immediate family—whether that relationship is provided by a family friend, relative, coach, teacher or formal program-based mentor. Youth who have less access to parental support in their lives, for example because of an absent or deceased parent, and those whose environments or personal characteristics or behaviors put them "at risk" for various negative outcomes have been widely viewed as having a special need for mentoring. In 2006, MENTOR: The National Mentoring Partnership estimated that 17.6 million youth in America could especially benefit from a mentoring relationship (MENTOR 2006).[9]

For youth within this group who are at higher risk for serious problems, such as academic failure, delinquency, violence or mental health difficulties, a mentor may be all the more crucial—and hard to come by in their everyday lives. Yet, as noted in Chapter 1, these youth may be more difficult for mentoring programs to serve, in part because their families may be too overwhelmed to seek out services and to support the youth's participation. Programs have to locate and connect with the families of such youth, interest them in mentoring (an intervention that invites an unknown adult into their lives, which many families may not want), and find and support volunteers who can meet the youth's needs.

In this chapter, we examine the concept of risk and the extent to which the programs in the study were already targeting higher-risk youth before the initiative. We then take a close look at the youth who were enrolled in the programs during the study period—the types of risks they faced and how they compare with other samples of youth nationally. (For a description of the mentors—their backgrounds, prior experiences and how the matches were created—see Appendix C.)

Defining Risk

This study was designed to examine how mentoring works for these higher-risk youth. However, defining this group of young people is more difficult than it might seem.

Risk is often thought of as existing on a continuum. Youth at the lower end of that continuum may be at risk for future problems. Yet, many—perhaps most—of them, with very little intervention, may actually do fine. In contrast, youth at the other end of the continuum may require intensive support to avoid serious problems down the road. Youth who are somewhere in the middle can be expected to experience problems at rates between the two extremes.

The continuum concept is useful shorthand, but in truth it masks the complexity of how risk factors actually play out in youth's lives. First, there is the question of "at risk for what?" Different background characteristics and events put youth at risk of negative outcomes in different areas (academics, health, delinquency, etc.).

Second, the youth's developmental stage helps determine which behaviors or attitudes and which features of his or her environment constitute risk. For example, skipping school or using drugs are not telling risk factors for very young children because they happen so rarely, whereas hitting or a lack of friends may be, depending on the outcome of interest.

Third, it may be important to consider both the *number* of risk factors a young person faces (Sameroff et al. 1998) as well as their *severity*.[10] Thus, a youth who lives in poverty is seen as at risk for dropping out of high school relative to her more affluent peers, whereas a young person who has been held back in school, has low academic achievement and has negative attitudes toward school may be viewed as being at even greater risk, or "high risk," for school dropout (see Janosz et al. 1997).[11] Yet, the same "*high*-risk" youth may not be at particularly high risk for, say, violence later in life. Moreover, one very serious risk factor (like homelessness, living in foster care or an arrest) has the potential to be just as powerful a predictor of negative outcomes as several less-serious risk factors. The extent to which risk factors are concentrated within one area of a youth's life, such

as within his family, or are present across multiple areas, such as peers, school *and* family, may also be important in determining his overall risk level (Gerard, Buehler 2004). Finally, risk factors that are present over time (such as chronic poverty; Korenman et al. 1995) are likely to have more impact than those that are episodic (for example, a parent's short-term unemployment)—a distinction that may be lost in assessments that focus on risk factors at one point in time.

To complicate matters even further, youth don't just bring risks to the table. Even those with highly challenging circumstances—such as those who have an incarcerated parent or who struggle at school or with mental health issues—often carry with them a host of environmental assets (loving and committed parents, for instance) and personal strengths and skills (such as positive aspirations for the future) that can help offset the risks they face.

It was not practical for this study to incorporate all of these facets of risk. We believed, however, based on available research, that two were crucial and thus focused on those: the number of risks youth experienced and the extent to which they experienced these risks in multiple domains of their lives. Although limited, prior research on mentoring and youth risk points to the need to distinguish between risks that stem from the youth's own behavior or characteristics (for example, academic challenges or mental health issues) and those that are more a reflection of the youth's environment and life circumstances (such as poverty or a high-crime neighborhood).[12]

In particular, as was noted in Chapter 1, meta-analyses of youth mentoring program evaluations (DuBois et al. 2002, 2011) suggest that youth's levels of "environmental" and "individual" risk, in combination, help shape the extent to which they benefit from mentoring. As we discuss in more detail below, this study builds on these findings by assessing the extent to which different profiles of individual and environmental risk characterize participating youth and then examining the implications of these profiles for the experiences of youth and their mentors as well as program effectiveness.

The Programs' Experience with Higher-Risk Youth

Seven programs participated in the evaluation, including five operated by Big Brothers Big Sisters of America (BBBSA) agencies. Distinct Big Brothers and Big Sisters organizations (the predecessors of today's combined Big Brothers Big Sisters organizations) were developed in the US more than a century ago to provide mentors for youth who were involved with the juvenile justice system (Baker, Maguire 2005). This mission eventually evolved into a focus on boys and girls from single-parent homes. Over the last 15 years, however, recognizing that many other groups of youth could benefit from program-based mentors, BBBSA agencies have worked to target youth with more diverse needs and backgrounds (for example, youth of color, youth with incarcerated family members, youth with a parent in the military, youth in foster care;

Finding and Enrolling Higher-Risk Youth

The programs in the study reached higher-risk youth in a number of ways, including collaborations with schools, partnerships with social service agencies, participation in community activities and events, word of mouth, and other media/communications strategies.[a] They reported that the most successful approaches were word of mouth (for example, families hearing about the program through other involved families, teachers or friends), partnerships with social service agencies and collaborations with schools. For the most part, these strategies were not new to the programs; most had used these approaches to reach higher-risk youth prior to the initiative.

The programs did not report significant challenges in finding higher-risk youth—especially the program that was already designed to serve such youth. However, staff did report a range of challenges in getting families of higher-risk youth interested and through the enrollment process. Some of the very issues that made youth "higher risk" (frequent moves, homelessness, foster care placement, for instance) also made it harder to get them enrolled. Working with families whose culture or past experience with social service providers may prevent "outsiders" from being trusted immediately created challenges as well.

[a] More than half of the programs reported using these approaches.

LaFleur, personal communication, Oct. 8, 2012). By 2009, BBBSA's strategic plan specifically addressed the organization's intent to serve youth facing more adversity in their lives. Most recently, BBBSA formed a task force focused on mentoring for youth who have been involved in the juvenile justice system, which interestingly means refocusing on a group of youth that its forerunner organizations originally set out to serve.

Today, most BBBSA agencies serve youth from ages 6 to 18. Across agencies (and often within them as well), youth come from a wide range of demographic backgrounds and have experienced a range of risk factors. Although all BBBSA agencies follow the same national guidelines for practice, they vary considerably in whom they serve, depending on their location, community partnerships and strategic direction. This is true even among the five BBBSA agencies in this study.

The study also involved two mentoring programs whose host organizations were not part of the BBBSA network. In the first—a university-based program—college volunteers meet with youth weekly at the youth's school (during the school day) and at the university. The program also invites youth's families to participate in monthly activities. Although the program does not specifically target "high-risk" youth, its feeder schools are located in high-poverty neighborhoods surrounding the university, and teachers refer youth who they believe need help and thus may be experiencing various challenges. The second program is run out of a larger agency that provides mentoring and a range of other services to multiple counties throughout Northwest Washington. The mentoring program targets youth in the community whose parents are in prison or have been in the corrections system. Thus, it is specifically designed to serve higher-risk youth.

How We Assessed Risk

To explore associations between youth's "risk profiles" and the relationships and benefits that are fostered through their program involvement, we collected a wide range of information on the number and types of challenges youth faced at the time of program enrollment. As a central component of these efforts, parents completed a 25-item screening tool as part of their baseline assessment (see Appendix D). To receive a preliminary status of "higher risk," youth needed to have at least one environmental risk factor (that is, challenges within the youth's life circumstances or relationships, such as housing instability, a family member with a drug or alcohol problem, or being bullied) and at least one individual risk factor (that is, challenges in the youth's behavior, academic functioning or mental health, such as gang involvement, poor school performance or depression). (See Risk Indicators Used in the Youth Risk Profiles text box on the next page.) In addition, the youth's total number of risk factors needed to equal or exceed four. Most of the programs in the study were asked to have at least 50 percent of enrolled youth fall into this higher-risk category; the two largest programs aimed for 75 percent. Youth who did not meet the higher-risk criteria were still able to participate in the program and the study. We set these targets to ensure that the study sample would have enough lower- and higher-risk youth to allow comparisons across groups.

The programs were quite successful at reaching youth who met or surpassed the higher-risk threshold, without significant efforts beyond their normal recruitment strategies. Almost two thirds (64 percent) of the youth enrolled in the study were deemed higher risk by this preliminary standard. Aside from the university-based program—in which less than half (38 percent) met the criteria for higher risk—there was little variation across programs, with a range of 64 to 77 percent of youth meeting the initial higher-risk criteria.

Although this preliminary standard was useful as a broad assessment of who the programs were enrolling, we wanted a more fine-grained, multidimensional measure of risk for the study's indepth analyses. Following the work of DuBois et al. (2002), we thus used the six risk domains noted in the Risk Indicators Used in the Youth Risk Profiles text box (which comprised 24 of the 25 items used in the preliminary risk assessment plus seven additional items) to develop four distinct risk groupings, or profiles: youth with relatively high individual

and high environmental risk (that is, the "highest-risk" youth in the sample), youth with relatively low individual but high environmental risk, youth with relatively high individual but low environmental risk and youth who were relatively low on both types of risk factors (the "lowest-risk" youth in the sample).

(See Appendix D for a description of how these groups were created.) These four groups are all "at risk" for various kinds of future problems, but, as discussed below, have very different characteristics.

Risk Indicators Used in the Youth Risk Profiles

We used parent responses to several survey items as indicators of youth "risk" at baseline. These are summarized below and described in more detail in Appendix D. All of the indicators listed below contributed to the risk profiles described in this chapter and used throughout the report in analyses comparing youth with different risk profiles. Twenty-four of the indicators were assessed with "yes/no" questions and were included (along with one additional item not listed here) in our preliminary screening tool. The remaining seven (starred below) were assessed using additional items that served to enhance the study's assessment of environmental risk.

Environmental Risk
Economic Adversity

- Youth's family lives in public housing*

- Parent believes family could be evicted or forced to leave home*

- Family has difficulty paying bills

- Youth's family lives in a neighborhood with a high level of gang or drug activity

- No working parent in the home*

- Youth's family is low income*

Family Risk/Stress

- Youth currently in foster care*

- Sibling (or youth) in foster care recently

- Parent/close relative incarcerated or often in trouble with police

- Family member with drug or alcohol problem

- Not living with both parents

- Moving two or more times in past year

- Recent parent separation*

- Frequent fights or arguments in home

- Recent death of/loss of contact with an adult youth knows well

- Recent homelessness

- Parent completing survey has less than a high school education*

Peer Difficulties

- No close friends

- Being bullied

Individual Risk
Academic Challenges

- Failing two or more classes

- Physical, emotional or mental condition interfering with grade-level school performance (for example, learning disability)

- Missing school three or more times a month

- Learning English as a second language

Problem Behavior

- Drug or alcohol use

- Two or more school suspensions in the last year

- Sent to juvenile detention or police contact

- Ran away from home

- Gang involvement

- Bullies others

Mental Health Concerns

- Exhibiting depressive symptoms

- Diagnosed with a mental health problem

Table 2.1
Parent-Reported Youth Demographics

Characteristic	Youth in the Study
Average age (years)	11.39
Male	53%
Race/Ethnicity	
African American	23%
Asian, Pacific Islander	5%
Hispanic	22%
Native American	5%
White	43%
Other race/ethnicity	3%

Note: The percentages reported in this table and elsewhere in this chapter are based on the 1,310 youth who were enrolled in the study. In other chapters, the sample sizes used for analysis differ from this number. One reason for these differences is that some youth in the control group subsequently received mentoring and thus contributed two observations to the analyses of program effects. The second primary reason is a lack of data for youth and parents who were not surveyed at follow-up and for mentors who did not complete a study survey.

Youth in the Study

The programs enrolled 1,310 youth in the study (including youth in both groups that had access to a mentor and the comparison group). These youth ranged from age 8 to 15 (97 percent were between 9 and 14), with an average age of a little over 11, and came from a variety of ethnic backgrounds—57 percent were ethnic minorities (see Table 2.1). Parents reported that nearly two thirds (66 percent) were from single-parent homes and that more than two fifths (43 percent) had family incomes lower than $20,000 per year (see Table 2.2 on page 21). Although very few of these youth were indicated to be in foster care (4 percent), a notable minority had experienced homelessness in the last five years, according to their parents (15 percent).

Are These Youth "High Risk"?

Nearly all youth in the study faced environmental adversity (one or more environmental risk factors listed in the Risk Indicators Used in the Youth Risk Profiles text box), and almost three quarters (71 percent) experienced one or more areas of personal vulnerability (one or more individual risk factors). Yet, relatively few youth were reported by parents to already be engaged in the kinds of behavior (for example, substance use, behaviors leading to involvement with police) that are frequently used to indicate "high-risk" status, perhaps in part because the youth in our sample were fairly young.[13] So, just how "at risk" were the youth in our study?

Our discussion earlier in this chapter underscores the difficulty in ascertaining the extent to which any group of youth are "high risk" (or, in the present context, "*higher* risk"). We did, however, try to put the risk status of our sample of youth in context, using three points of comparison—the average child in the US, a definition of high socio-demographic risk developed by Child Trends (Moore 2006; Moore et al. 2006) and the average youth served by BBBSA agencies at a national level (see Table 2.2).[14] We found that the youth in our study appear to be much more likely to face a number of risk factors than the average child in the US. For example, youth in this sample were more than twice as likely to live in extreme poverty (and more than two and a half times as likely to receive free or reduced-price lunch), almost twice as likely to come from single-parent homes, almost four times as likely to report serious signs of depression and several times as likely to live in foster care. The youth were much more similar to youth served by BBBSA agencies nationally. Compared with youth in the BBBSA sample, youth

Table 2.2
Challenges Faced by Youth in the Study and Nationally

Challenge	Youth in This Study (Percent of Parents Reporting This Youth Characteristic or Need)	Percent of Youth Nationally	Percent of Youth Served by BBBSA[a]
Live in single-parent household	66%	34%[b]	71%
Free or reduced-price lunch	77%	28%[c]	61%
Annual income below $20,000 (or food stamp receipt)	43%[d]	22%[e]	
Live in "extreme" poverty (income of $10,000 or less)	23%	10%[f]	
Experienced homelessness in the last five years	15%	2%[g]	
Close family member incarcerated or having frequent problems with law	24%[h]	4%[i]	23%
Suspended from school two or more times in the last year	12%	7%[j]	
Sent to juvenile hall or had contact with the police in the last year	6%	3%[k]	
In foster care	4%	0.5%[g]	1%
Used drugs, alcohol	4%		
Serious signs of depression	22%[l]	6%	
Environmental difficulties[m]			
Economic adversity	83%		
Family risk/stress	91%		
Peer difficulties	52%		
At least one risk factor in each of the three environmental domains	41%		
At least one environmental difficulty	99%		
Individual difficulties[m]			
Academic challenges	53%		
Mental health concerns	48%		
Problem behavior	23%		
At least one risk factor in each of the three individual domains	14%		
At least one individual difficulty	71%		
Either **individual or environmental difficulties**	99%		
Both **individual and environmental difficulties**	71%		
Designated as "high" on both	26%		
Designated as "high" on at least one	64%		
Designated as "low" on both	36%		

Table 2.2
Challenges Faced by Youth in the Study and Nationally, *continued*

Note: Each risk area had different numbers of indicators (that is, questions to which a parent could respond "yes" to be indicated as facing that type of difficulty)—see Risk Indicators Used in the Youth Risk Profiles text box for more details and Appendix D for the specific questions that contributed to each category.

[a] BBBSA statistics from the Agency Information Management (AIM) system reflecting all 9- to 14-year-olds served in BBBSA community-based programs between July 1, 2011 and June 30, 2012, in agencies using the AIM system (more than 65,000 youth; Wheeler, personal communication, Nov. 14, 2012, BBBSA statistics for 2011).

[b] Statistics for 2010 (*Kids Count Data Book* 2012).

[c] Statistics for 2011 (*State of America's Children*, Children's Defense Fund 2012).

[d] Income was derived from a question asking parents to select the range in which their income fell; 5 percent of parents left this question blank. For those parents, we used parent reports of food stamp receipt as a proxy for having an income below $20,000.

[e] Percent of youth below the poverty line in 2010 (*State of America's Children*, Children's Defense Fund 2012).

[f] Percent of youth living in families with incomes 50 percent or less of the poverty level in 2010 (*State of America's Children*, Children's Defense Fund 2012).

[g] Statistics for 2010 (*State of America's Children*, Children's Defense Fund 2012).

[h] Percent of youth in our sample whose parents reported that a parent or close family member is incarcerated or frequently in trouble with the police.

[i] Percent of youth with an incarcerated parent (The Pew Charitable Trusts 2010).

[j] Percent of youth suspended one or more times in 2009 (*State of America's Children*, Children's Defense Fund 2012).

[k] Percent of youth who were arrested as juveniles in 2009 (*State of America's Children*, Children's Defense Fund 2012).

[l] This is a youth-reported measure. The percentages refer to youth scoring 11 or higher on the Short Mood and Feelings Questionnaire (SMFQ; Angold et al. 1995), which measures child and adolescent depression. The comparison percentage is based on a large community (not national) sample (Angold et al. 2002). Angold et al. (2002) refer to a score of 11 or higher as an "extreme" score on this measure. Of note, 37 percent of our sample scored 8 or higher—a cutoff that is highly predictive of depression status (Angold et al. 1995). We use the more extreme cutoff in defining "serious signs of depression" to enable us to compare our sample with another large sample of youth.

[m] See Risk Indicators Used in the Youth Risk Profiles text box in this chapter for definitions of each of these categories.

in this study were slightly less likely to live in a single-parent home, but they were more likely to receive free or reduced-price lunch and to live in foster care.

In addition, Child Trends defined a youth as being at "high socio-demographic risk" if he or she faces three or more of the following five risk factors: poverty, living in a single-parent family, both parents (or the child's only parent) having a low level of education (that is, less than high school graduation or GED), a large family (three or more children) and a family that is not able to own or buy a home.[15] While 18 percent of a national sample of youth ages 0 to 17 fit this definition of high risk, one quarter (25 percent) of the youth in our study did. And whereas 38 percent of the national sample had none of these risk factors, this was true for only 15 percent of our sample.

Therefore, although a substantial proportion of the youth in this study meets the Child Trends definition of high risk, most of them do not. Still, the youth in the study are clearly higher risk, on average, than the typical American child. The number of risk factors they faced across multiple domains (nearly three quarters were experiencing at least

one environmental and one individual risk factor) suggests that, without intervention, many could encounter serious problems down the road.

Examining the characteristics of youth within each of the four risk profiles provides a more detailed description of youth's experiences and shows how very disparate their needs were across the four risk groups. (See Figure 2.1 on the next page for more detail.) The youth in the "lowest-risk" group (that is, those who were relatively low on both environmental and individual risk; about 36 percent of the youth in our sample) were relatively unlikely to experience most of the risk factors we depict in Figure 2.1. Less than 5 percent had exhibited one or more of the problem behaviors noted, and although 58 percent lived in single-parent homes, only about a third lived in poverty. Youth high on only individual risk (about 14 percent of youth) were relatively more likely to have had frequent school suspensions and to have had police contact in the last year, and almost a quarter showed serious signs of depression at baseline. But, on average, their home environments appeared to be the most stable of all four groups, with a little over half living in single-parent homes but under a quarter living in poverty. This is in stark contrast to

Figure 2.1
Key Characteristics of the Four Risk Groups

	Low Individual Risk	High Individual Risk

Low Environmental Risk — Low Individual Risk: Serious signs of depression 18%, Police contact 1%, Frequent suspensions 4%, Close family member incarcerated or often in trouble with police 15%, Homelessness 6%, Single-parent home 58%, Poverty 32%. High Individual Risk: 24%, 9%, 19%, 18%, 6%, 52%, 23%.

High Environmental Risk — Low Individual Risk: 18%, 1%, 4%, 29%, 21%, 80%, 60%. High Individual Risk: 32%, 15%, 28%, 35%, 26%, 71%, 54%.

Legend:
- Serious signs of depression
- Police contact
- Frequent suspensions
- Close family member incarcerated or often in trouble with police
- Homelessness
- Single-parent home
- Poverty

those youth who were relatively high on only environmental risk (about 23 percent of youth) and appeared to be living in the most challenging environments. The vast majority of these youth (80 percent) lived in single-parent homes, and 60 percent lived in poverty. About a fifth had also experienced homelessness in the last five years, and 29 percent had a parent or other close family member who was incarcerated or frequently in trouble with the law. Finally, our highest-risk youth (those high on both environmental and individual risk; about 26 percent of study participants) were relatively likely to have experienced several of these challenges and would certainly, by most accounts, be characterized as "high risk." Almost a third showed serious signs of depression at baseline; 15 percent had had recent contact with the police; and 28 percent had frequent school suspensions. Their home lives were also troubled, with 26 percent experiencing homelessness in the last five years, 71 percent living in a single-parent home, 35 percent having a parent or close family member incarcerated or frequently in trouble with the law, and more than half (54 percent) living in poverty. Clearly, youth in all four groups were experiencing challenges in their lives—several statistics for even the lowest-risk group were higher than those for youth at a national level. But, life experiences differed greatly across the groups, which likely created very different contexts for the development of mentoring relationships.

In addition to parents' reports of youth's characteristics and experiences, we were also interested in the needs mentors saw in their mentee and how they aligned with those reported by parents. To explore this issue, we asked mentors what *they* perceived were the areas of need or challenge for their mentee.[16] We could not directly compare mentor and parent responses for individual children, because the two surveys asked questions about youth risk in different ways. In general, however, mentors reported somewhat fewer difficulties than those reported by parents. For example, relatively few mentors reported health or mental health concerns (15 percent), and somewhat less than a third reported poverty as a challenge for their mentee (29 percent)—the latter, a risk factor that would seem to be readily apparent to mentors who visited the youth's home and community. Perhaps mentors bring with them different understandings of poverty than the income-based definition used when considering data from the

parent survey. Alternatively, some may have simply interpreted our question as asking about problems they could help youth with in their meetings—and because poverty is not a problem that could be "solved," they may have ignored it in their responses. The discrepancy between mentor and parent reports, however, does raise questions about whether some mentors fully appreciated the level of need of the youth with whom they were matched—a theme we return to in the next chapter.

Summary

The youth who were recruited for this study experienced a wide range of significant challenges in their lives. As a group, they were significantly "higher risk" than the average American youth, but only about a quarter met Child Trend's definition of high socio-demographic risk, and few had engaged in the kinds of behaviors (for example, substance use, crime) that are often used to determine "high-risk" status—perhaps in part because the youth in our sample were fairly young. For younger children, multiple school suspensions and bullying may be just as telling, as they predict more serious problem behaviors down the road (almost a quarter of the sample had engaged in these types of problem behaviors—bullying, etc.—which are more relevant for younger children). Interestingly, programs reported that higher-risk youth were generally not difficult to *find* in their communities, but they were sometimes difficult to enroll.

It is important to note that there was a great deal of variation in the levels and types of risk faced by youth in our study sample. Although most seemed to fall somewhere in the middle of a broadly defined risk continuum, there were clearly youth at both ends of the spectrum. Indeed, it is this variation that allowed us to examine how youth with very different risk profiles experienced and benefited from mentoring. Throughout the report, we compare the four risk subgroups described in this chapter to understand how the mentoring programs worked for different kinds of youth.

In the next chapter we look specifically at the relationships that were forged between youth and their mentors. We focus on the quality and length of the relationships, as well as challenges that emerged for the mentors.

What Kinds of Relationships Did Youth Experience?

Chapter III

The relationships that develop between youth and their mentors are the central route through which mentoring is thought to benefit youth (Rhodes 2005). By developing trusting, caring relationships with adults, youth may begin to see themselves and others in a more positive light, develop cognitive and social skills that they can transfer to other important relationships and envision their futures with greater optimism and clarity (Rhodes 2005). In support of this view, studies suggest that the quality of the mentoring relationship matters: Stronger and longer mentoring relationships tend to be linked with more favorable outcomes for youth (Grossman, Rhodes 2002; Herrera et al. 2007; Rhodes, DuBois 2006).

Success in creating strong, long-lasting and potentially life-changing mentoring relationships is certainly not a given in any mentoring program. Even in Big Brothers Big Sisters—the most well-established and largest mentoring program in the country—less than two thirds (64 percent in 2011) of the community-based matches sustain their relationship up to or beyond the mentor's 12-month commitment (Valentino, personal communication, Oct. 29, 2012, BBBSA statistics for 2011).

Youth characteristics and needs also play an important role in relationship development. Higher-risk youth, in particular, can present numerous challenges to mentors. Youth may come to relationships with social or behavioral difficulties that mentors are ill equipped to handle. And many higher-risk youth have experienced unhealthy relationships or the loss of adults who were important to them. As a result, they may be untrusting of adults and reluctant to invest in a new relationship, or may test mentors to see if they are really committed (Ahrens et al. 2011). Frequent moves, health issues and unstable family lives (for example, changing guardianship, parent incarceration, homelessness, lack of regular phone access) also may make it very difficult to sustain mentoring relationships with higher-risk youth. In line with these considerations, in P/PV's National Faith-Based Initiative (NFBI) demonstration project, older youth and those who

What Makes for a Strong Mentoring Relationship?

Relationship strength has typically been assessed using measures that focus on how close or emotionally "connected" mentors and mentees feel toward one another (for example, Herrera et al. 2007; Parra et al. 2002). Recent work (Karcher, Nakkula 2010; Nakkula, Harris 2005, in press), however, has emphasized a more multidimensional view of relationship quality. These frameworks call attention to both "relational" and "instrumental" dimensions of mentoring relationships. Relational qualities involve how the youth and mentor experience or feel about their relationship, whereas instrumental qualities typically involve how the mentor may be helping the young person to pursue goals or personal growth.

Relational qualities include feelings of closeness in the relationship as well as the extent to which the relationship is "youth centered." This latter feature of relationships was described in Morrow and Styles' (1995) seminal research on mentoring relationship development in community-based programs. In youth-centered relationships (which were more likely to be sustained over time), the mentor actively considered the youth's interests and was not preoccupied with achieving changes in the youth's behavior at the expense of developing a positive connection, especially in the early stages of the match.

At the same time, research on the instrumental dimension of mentoring relationships has indicated that engaging in purposeful activities together can help to deepen the relationship and make it more meaningful (Hamilton, Hamilton 1990). Similarly, research suggests that introducing goals can help focus match activities on interests or areas of concern that are important for the young person's well-being or development (Balcazar, Keys in press).

Although on the surface these two aspects of mentoring relationships (that is, their relational and instrumental qualities) may appear to be in conflict, they need not be (Nakkula, Harris 2005, in press). In fact, there are good reasons to believe they can be complementary or synergistic, such as when mentors work closely with youth to identify goals that are of particular interest to them (see Keller 2005).

Analyses in the current study support the idea that youth centeredness and a growth/goal orientation are distinct but associated dimensions of mentoring relationship quality. That is, youth who rated their mentor higher in youth centeredness tended to also report more of a growth/goal focus in the relationship (see Appendix E for details).

Designations of the Four Risk Profiles

Throughout this report, we characterize the four risk profiles as follows:

- **High Env** = The group relatively high only on environmental risk

- **High Ind** = The group relatively high only on individual risk

- **High/High** or **"Highest Risk"** = The group relatively high on both types of risk

- **Low/Low** or **"Lowest Risk"** = The group relatively low on both types of risk

In addition, when referring to *both* groups that are relatively low on environmental risk, we use the term "low environmental" or "**Low Env.**" When referring to *both* groups that are relatively low on individual risk, we use the term "low individual" or "**Low Ind.**"

---

were higher risk (that is, had been arrested previously) were less likely to sustain relationships with their mentor (Bauldry, Hartman 2004).[17] Grossman and Rhodes (2002) similarly found that matches with older youth; with youth who had sustained emotional, sexual or physical abuse (environmental risk); and with youth who had been referred for psychological or educational programs (individual risk) were shorter than those with youth who didn't share those characteristics.

In this chapter we describe the mentoring relationships of youth in both mentored groups (that is, all youth who had access to a mentor through both the random assignment and quasi-experimental components of the study), using data recorded on each match by programs to describe the frequency and duration of their meetings and the length of the relationships that developed. We also use mentor[18] and youth reports to describe the quality of the relationships, the activities matches engaged in, why many of these relationships ended prior to the mentor's stated commitment and the challenges mentors encountered in developing them. Throughout the chapter, we compare youth with different risk profiles (the four subgroups introduced in Chapter 2) to shed light on how risk may have played a role in shaping the length or strength of their mentoring relationships.[19]

Match Meetings

Mentors in the initiative were asked to meet with their mentee at least twice a month. Mentors who were in a match that received program enhancements in one of the two largest programs (that is, those involved in the randomized trial portion of the evaluation),[20] or in one of two additional programs, were asked to meet with their mentee at least three times a month.[21] During the study period, across all programs, matches got together an average of 2.0 times per month. Program records indicated that only about half (46 percent) of matches were able to meet the expectation for meeting twice monthly at least 70 percent of the time; a small minority (15 percent) met at least three times monthly 70 percent or more of the time (see Table 3.1 on the next page). Matches met for about three hours per meeting, with an average of 70 hours of mentoring across the entire match, as measured at follow-up. In addition, program records show that about three quarters (76 percent) of matches also communicated through other means (that is, "non-face-to-face" communication, such as email or phone conversations). On average, this type of communication occurred in about two fifths (42 percent) of the months that matches were together. The frequency or duration of meetings as well as the presence or frequency of non-face-to-face communication did not differ for youth with different types of risk.[22]

Cancellation of match meetings was not uncommon. Overall, slightly more than half of the mentors (55 percent) reported that they had canceled at least one scheduled meeting. A larger proportion (72 percent) reported that youth (or presumably their parents) had canceled at least one meeting. For 18 percent of matches, youth were reported to have canceled meetings fairly frequently—every month or two or more often (only 3 percent of mentors reported that they canceled meetings this often). Such frequent youth cancellations were particularly common for mentors who were matched with youth who were relatively high on environmental risk[23] and were associated with matches meeting somewhat less often, on average, each month.

Mentors often struggled to fit match commitments into their already busy schedules—so, having meetings canceled, sometimes with no warning, was very frustrating for many mentors (a theme we discuss later in the chapter).

Table 3.1
Match Meetings

Match Meetings	Overall	Association with Risk[a]
Face-to-face meetings[b]		
Met 2+ times per month 70% or more of their months together	46%	None
Met 3+ times per month 70% or more of their months together	15%	None
Average number of meetings per month	2.0	None
Average number of hours per meeting	2.8	None
Average number of hours met per month	5.8	None
Total hours of meetings throughout entire match prior to follow-up	69.8	None
Non-face-to-face meetings[b]		
Had non-face-to-face communication during match	76%	None
Average percent of months with non-face-to-face communication	42%	None
Missed meetings[c]		
Percent who missed meetings due to youth cancellation every month or two or more often	18%	High/High & High Env > Low /Low
Percent who missed meetings due to mentor cancellation every month or two or more often	3%	NT[d]

[a] Tests of associations with risk controlled for youth gender, age and race/ethnicity as well as which mentoring program served the match.

[b] Measures are based on records maintained by programs.

[c] Measures are based on mentor reports.

[d] Differences across risk groups were not tested (NT) because the distribution of responses on the measure was outside the 10- to 90-percent range (for example, one group encompassed 5 percent and the other, 95 percent) and thus resulted in a very small group in the less frequent response category.

Activities

Community-based matches engage in a wide range of activities. Mentors are given general parameters they must follow; these vary by agency, even within BBBSA programs, but may include, for example, no visits to the mentor's home (or visits only after a certain amount of time into the relationship), no expensive activities and no overnight outings. As long as matches follow those broad parameters, they are free to engage in activities of their choosing.

To understand more about how matches spent their time together, we asked mentors about five broad categories of activities they might have engaged in with their mentee (see Table 3.2 on the next page).

Almost all mentors reported spending large amounts of their time "having fun" (for example, "making time to goof around, laugh and have light-hearted fun" with their mentee) and talking about a wide range of issues. About a quarter (26 percent) also reported spending a good deal of time on enrichment activities (for instance, cultural events, service activities, taking a class together), or working toward character or behavior change (for example, teaching mentees to manage their behavior or develop social skills). Fewer (17 percent) reported spending significant amounts of time in structured or program-based activities, such as program-wide events, spending time with other matches or meeting with other professionals on behalf of the youth.[24]

Table 3.2
Mentor-Reported Match Activities

Activities	Percent Reporting Spending A Lot of Match Time Engaged in the Activity[a]	Association with Risk[b]
Fun	95%	NT[c]
Talking	84%	None
Character/behavior change	53%	High/High > Low Ind
Enrichment activities	26%	None
Structured/program activities[d]	17%	None

Note: Each of the activity categories represents one or more single-item questions.

[a] The mentor either had an average response at or above 2 ("some of our time together") on the full combined scale or responded to at least one item within the scale with a 3 ("a lot of our time together") or 4 ("most of our time together").

[b] Tests of associations with risk controlled for youth gender, age and race/ethnicity as well as which mentoring program served the match.

[c] Differences across risk groups were not tested (NT) because the distribution of responses on the measure was outside the 10- to 90-percent range (for example, one group encompassed 5 percent and the other, 95 percent) and thus resulted in a very small group in the less frequent response category.

[d] The university-based program was excluded from analyses of this category because the program was very structured.

Mentors also reported how often they—versus their mentee or someone else—made decisions about how they would spend their time together (see Table 3.3 on the next page). Youth enjoy and benefit from having a voice in the activities they engage in, and, as noted earlier, studies suggest that when youth contribute to decision-making they have more successful mentoring relationships (Morrow, Styles 1995). A little over two thirds (69 percent) of mentors reported this kind of joint decision-making. About 5 percent said their mentee usually decided how they spent their time together, and 4 percent said the program typically outlined their activities (these youth were almost exclusively from the university-based program,[25] in which matches engaged in structured activities on the university campus). Only about 1 percent indicated that the youth's parents made these decisions. An additional 22 percent of mentors reported that they typically decided themselves.

We were also interested in the broad focus, or goals, of the mentoring relationships. We hypothesized that mentors who were matched with youth who had more serious needs might focus on helping the youth improve in those particular areas—either because the mentor saw a need or because the youth's parents, program staff or the youth themselves highlighted that need.

Two thirds of mentors reported that their main goal, or focus area, for the match was to be a friend to their mentee—a goal that Morrow and Styles (1995) found was linked with longer-lasting relationships that youth regarded as more satisfying. This likely reflects the fact that most mentors in this study participated in BBBS programs, which typically have a purely friendship-based approach to their work with youth (as opposed to focusing on targeted goals). Close to a third (29 percent) noted that helping the youth improve in some way (for example, their self-esteem, behavior or academics) was their central goal or focus.

When asked which factors led them to choose their specific focus, a majority of mentors (61 percent) reported that they had noticed this as a key need for the mentee.[26] Nearly half (45 percent) reported choosing the area of focus based on what they believed youth of this age typically need. About 1 in 5 (19 percent) indicated that program staff helped them determine the goal; the same proportion indicated that their mentee's parents helped to determine the match's focus. Notably, only a small proportion of mentors—about 12 percent—reported that they and their mentee jointly, or their mentee alone, determined the goal or area of focus. It is difficult to say why this proportion was so much lower than the proportion of mentors who

Table 3.3
Mentor-Reported Match Decision-Making and Goals

Decision-Making and Goals	Percent Agreeing	Association with Risk[a]
Who usually decided how match spent its time?[b]		
Mentor and youth together	69%	None
Mentor	22%	—[c]
Youth	5%	—
Program staff	4%	—
Youth's parent(s)	1%	—
Mentor's most important goal/focus for the match?		
Friendship	67%	None
Improvement (for example, academic, behavioral, self-esteem)	29%	None
Setting personal goals	2%	NT[d]
Exposing youth to new opportunities	2%	NT
How was the goal/focus for the match determined?[e]		
Program staff	19%	—
Youth's parent(s)	19%	—
Mentor noticed as a key need for the mentee	61%	—
Mentor felt this was a need for youth of this age	45%	—
Mentee asked for help in this area or mentee and mentor decided together	12%	High/High < Low Env

[a] Tests of associations with risk controlled for youth gender, age and race/ethnicity as well as which mentoring program served the match.

[b] Differences were tested for decision-making for only the first category (mentor and youth together) because we were most interested in whether joint decision-making was affected by youth's risk profile.

[c] Differences across risk groups were not tested because we elected to test a limited set of hypotheses for this variable.

[d] Differences across risk groups were not tested (NT) because the distribution of responses on the measure was outside the 10- to 90-percent range (for example, one group encompassed 5 percent and the other, 95 percent) and thus resulted in a very small group in the less frequent response category.

[e] Mentors could choose more than one factor as determining the match's goal/focus. Differences were tested for goal determination for only the last category, because we were most interested in whether involving youth in these decisions (that is, whether youth contributed to the goal/focus) was affected by youth's risk profile.

reported engaging in joint decision-making with their mentee. Perhaps many volunteers simply felt that determining the underlying focus of the match was their task, given that these youth were, on average, fairly young. Programs may also have stressed the importance of "friendship" goals, such that volunteers readily saw this as the appropriate focus for the relationship.

Generally, matches did not differ in their activity choice, goals or decision-making based on mentee's risk profile. However, notably, engaging frequently in character/behavior change activities was reported most often by mentors matched with youth who were high on both types of risk.[27] This same group of youth (high on both types of risk) was the *least* likely to have input about the focus for the relationship.[28]

Relationship Quality

In addition to asking mentors about the activities they engaged in with youth, we also wanted to understand more about the quality of these interactions. We assessed relationship quality by asking the youth and the mentor several questions about their perceptions of the relationship. Mentors were asked how close they felt to their mentee, and youth were asked several sets of questions about how they felt when with the mentor (that is, how close they felt to him or her, how much their activities were centered on youth's preferences and the extent to which their relationship focused on helping youth achieve specific goals or personal growth).

Overall, mentors and youth reported fairly positive relationships (see Table 3.4). Almost three fifths of mentors agreed or strongly agreed to feeling "close" to their mentee. A similar proportion of youth reported feeling "very close" to their mentor. (Proportions for both were stronger—71 percent— when considering only mentors and youth in ongoing matches—that is, those that had not already ended by the time follow-up surveys were administered.) These rates for closeness are very similar to those reported in a study of BBBSA school-based mentoring (Herrera et al. 2007).[29]

Youth strongly agreed that the relationships were centered on their preferences (for instance, about activities). The average rating by youth on this scale (3.73) compares favorably with the average rating by youth on the same scale in the P/PV impact study of the BBBSA community-based mentoring program (3.53; Saito 2001). To a somewhat lesser extent, youth agreed that their relationships were focused on their personal growth or goals.

Reports of relationship quality generally did not differ based on youth's risk profile (see Table 3.4)—that is, youth and mentors were similarly positive regardless of youth's risk profile.[30] We also looked at these data using a criterion for match quality that required favorable youth reports of closeness, youth centeredness *and* growth (in particular, youth were required to score a 3.0 out of 4.0 or higher on all of these scales). Nearly three quarters of mentoring relationships (71 percent) met this criterion.[31] The rate at which such relationships were reported did not differ based on the youth's risk profile.[32] Several mentors commented on how positive their relationship was:[33]

> *My match and I have lots of fun, trust each other and share a lot with one another. I have enjoyed her company and am thankful for the opportunity to get to know her and to regularly get out in the community and 'play.' And I feel good that she enjoys hanging out with me and sees me as a friend....*

Table 3.4
Mentoring Relationship Quality

Relationship Quality[a]	All	Association with Risk[b]
Mentor-reported "close" relationships[c]	59%	None
Youth-reported "very close" relationships[d]	61%	None
Relationship was growth/goal focused[e]	3.27	High/High > Low/Low
Relationship was youth centered[e]	3.73	None

[a] All scales are youth reported unless otherwise indicated.

[b] Tests of associations with risk controlled for youth gender, age and race/ethnicity, which mentoring program served the match and, in the case of youth-reported measures, mode of survey administration (phone or in-person).

[c] Percent who "agreed" or "strongly agreed" (that is, responded with a 4 or 5 on a 5-point scale) to the statement, "I feel close with my mentee."

[d] Percent who responded "very close" (that is, responded with a 4 on a 4-point scale) to the question, "How close do you feel to your mentor?"

[e] The average score on the measure is reported, where 1 = "Not at all true," 2 = "Not very true," 3 = "Sort of true" and 4 = "Very true."

Others noted that they saw the mentoring relationship as an enduring part of their life, not simply a time-limited experience:

> *Teyara[34] and I have talked about being matched for a very long time. It's gone beyond a match— it's a friendship, and there's love there.*

As we discuss later in this chapter, however, many mentors reported significant challenges in their relationship. In many cases, even good relationships took time and patience to develop:

> *I've been matched for under a year, and she needed more time than average to trust me. Because of that, our activities have been limited, but I can see progress and will be patient for her to take the next step.*

In sum, the relationships in this study, as reported by youth and mentors, were generally fairly positive and were comparable in closeness and youth centeredness to those reported in past mentoring evaluations. It is notable—and encouraging—that relationship quality was similarly strong for youth with varying risk profiles. In the next section, we examine whether the *length* of the mentoring relationship differed on the basis of the youth's risk factors.

Match Length and Closures

While mentors in this initiative were asked to commit to at least 12 months (as is true in many community-based programs), the goal is generally to sustain matches for as long as possible—reflecting research that suggests longer matches yield stronger benefits for youth (Grossman, Rhodes 2002). In fact, one of the enhancements that were implemented by the two largest programs for matches in the randomized trial part of the evaluation (and by one additional program) was to ask mentors to commit to continuing their matches for 18 months or more, if possible.

When a match closes, and youth are interested in being paired with another mentor, many programs currently have policies in place to try to "rematch" youth with another mentor. For example, in 2010, for all youth exiting a community-based match in BBBSA programs, 67 percent were completing their first match, 25 percent were completing their

second match, and 8 percent were completing their third or higher match (Wheeler, personal communication, Oct. 29, 2012, BBBSA statistics for 2010). The hope is that youth who are rematched will receive "more" mentoring and benefit more than if they had not continued meeting with a mentor. The programs in the study similarly tried to rematch youth as needed.

By the time of the follow-up survey, about 13 months after baseline, almost half the youth in the combined treatment and participant group (48 percent) had experienced at least one match closure; about 60 percent were in ongoing matches. At that time: 4 percent of the youth in the combined group had never been matched; 85 percent had been involved in only one match; and 11 percent in two.[35] Grossman et al. (2012) found that youth who were rematched in the BBBSA school-based mentoring program actually experienced setbacks relative to youth in the control group. We wanted to test whether this was also true in this community-based sample, as it is possible that their findings were attributable to something unique about the school setting. In addition, we wanted to test whether previous links between match length and outcomes (for example, Grossman, Rhodes 2002) held true in this sample. We did not find evidence that longer matches or rematching in general were associated with differential outcomes for youth. However, there was some evidence that rematching was implicated in poorer outcomes to the extent that rematched youth were in mentoring relationships for relatively greater amounts of time (and these matches actually evidenced a pattern of *weaker* benefits). (See Match Length, Rematching and Youth Outcomes text box on the next page and Appendix F for more detail.)

At the time they took their follow-up survey, youth had experienced an average of 9.58 months of mentoring across all of their matches,[36,37] with an average of 8.95 months for youth's first match. Approximately 15 percent of first matches closed in less than three months, an additional 15 percent closed between three and six months, and 17 percent lasted at least six months, but ended earlier than the one-year mark. The remaining 53 percent surpassed one year in duration.[38,39] Importantly, neither the length of time youth received mentoring nor the rate of early closures (that is, matches that closed before 12 months) varied on the basis of the

Match Length, Rematching and Youth Outcomes

Previous research (for example, Grossman, Rhodes 2002) has suggested that longer mentoring relationships facilitate stronger outcomes for youth. We conducted analyses to see whether this pattern was also evident in this study (see Appendix F for details). Among youth who had not been rematched by follow-up, benefits did not appear to vary depending on the length of their match (that is, those with longer matches did not seem to benefit more than those with shorter matches). Interestingly, however, among youth who *had* been rematched, program effects for certain outcomes appeared to dissipate with greater length of time mentored. In these instances, rematched youth with the most total time mentored tended to have the poorest outcomes relative to youth from the comparison group. These findings caution against using rematching to achieve greater "doses" of mentoring for youth whose original matches have ended. It could be, for example, that such youth are not good candidates for mentoring (at least as provided by the program or at the youth's specific stage of development), such that simply adding more mentoring is of little value and may even be counterproductive. Risk for harm may be especially great when one or more of the relationships involved "drag on" despite indications of not being helpful or of high quality.

youth's risk status. However, among youth who had experienced a closure, those who were high on individual but low on environmental risk were the least likely to be rematched.[40] It could be that parents of such youth felt they had enough support in their home environment to overcome the child's personal or behavioral challenges, without continued intervention. Because these matches were more likely than those in other risk groups to be closed at the program's request (as discussed below), it could also be that the program simply felt that another match would not be productive for these youth or that their behaviors might make them difficult to rematch.

Of the mentors who responded to our survey,[41] 31 percent were reporting on a relationship that had already ended, and 4 percent on a relationship that would end soon. Three fifths (61 percent) of these mentors reported that they had initiated the closure (or impending closure) themselves, whereas only 23 percent said the closure had been initiated by the

youth, 19 percent by the program and 17 percent by the youth's parents (see Table 3.5 on the next page). Thus, it appears that it was typically the mentor who ultimately determined the match's length. Who was reported to have initiated the match's end generally did not differ by youth's risk; however, as noted, matches with youth high on only individual risk were most likely to be closed by the program.[42]

Mentors reported a variety of reasons for match endings, and many times these reasons were related to the types of risk a mentee was facing. About a third of the mentors reported that the match closed because the youth did not have enough interest in the match. This was particularly common among mentors matched with youth who were low on both types of risk.[43] Similarly, 17 percent of mentors reported that the youth did not seem to need a mentor—a reason that was most often cited when youth were low on environmental risk (perhaps because they appeared to have enough support in their homes) and least often cited when youth were high on both types of risk.[44] These situations were difficult for mentors—many of whom were already working hard to make time for their matches:

> *It has been hard for Selma and I to get together at times, she is either not very interested in getting together or is very busy…. Because of this, there have been times I wonder if she really needs a mentor; her schedule seems very full with supportive people and environments. Am I needed? At times it seemed that she felt like meeting me was a chore she needed to do, and after a few weeks, I felt the same way.*

A little more than a quarter of mentors (28 percent) reported that the match closed because the program no longer fit into their schedules, and 13 percent indicated that their mentee was moving. Another 17 percent reported that their match closed because the youth's family was not supportive, and 14 percent cited differences in interests or personalities. This last reason was most common in matches with youth who were high on both types of risk.[45]

Surprisingly (given the risk profile of many of the youth involved in the study), only 10 percent indicated that their match had closed or was closing because the youth's needs were too severe (a reason cited, as would be expected, most often when

Table 3.5
Match Length and Closures

Match Length and Closures	Overall	Association with Risk[a]
Average length of first match by follow-up	8.95 months	None
Average duration of all matches combined by follow-up	9.58 months	None
Percent of those with a closure who were rematched	24%	High Ind < all others
Mentor-reported reasons for closure[b]		
Not enough youth interest	33%	Low/Low > High Ind & High Env
No longer fits in mentor's schedule	28%	None
Youth did not seem to need a mentor	17%	Low Env > High/High
Youth's family was not supportive	17%	None
Difference in interests or personalities	14%	High/High > Low Ind
Mentee is moving	13%	None
Youth's needs were too severe	10%	High/High > Low/Low
Not enough program staff support	3%	NT[c]
Difficulty bridging cultural differences	<1%	NT
Percent of first matches closed before 12 months	47%	None
Initiation of closure[b]		
Mentor	61%	None
Youth	23%	None
Program staff	19%	High Ind > Low/Low
Mentee's parents	17%	None

[a] Tests of associations with risk controlled for youth gender, age and race/ethnicity as well as which mentoring program served the match.

[b] Limited to mentors who reported in the mentor survey that their match had closed or was about to close. Mentors could choose more than one option for this question.

[c] Differences across risk groups were not tested (NT) because the distribution of responses on the measure was outside the 10- to 90-percent range (for example, one group encompassed 5 percent and the other, 95 percent) and thus resulted in a very small group in the less frequent response category.

youth were high on both types of risk and least often when youth were low on both types of risk).[46] It is notable that more mentors reported that their match closed because youth didn't seem to need a mentor than because youth's needs were too severe.

About a quarter (27 percent) of mentors who were in closed or closing matches said they were "very disappointed" about the match ending, and about half (54 percent) said they were "somewhat disappointed." These reports did not differ by risk status,

suggesting that, in general, mentors did want to make their relationship work.

About two fifths of the mentors who'd experienced a closure (42 percent) said they would be matched again with another child. Interestingly, mentors matched with the highest-risk youth were the most likely to say they *would* be matched again.[47] Thus, mentors who had worked with youth who were most needy were the most likely to want to continue mentoring.

Mentor-Reported Challenges

Renee Spencer's qualitative study, examining the challenges inherent in sustaining matches (Spencer 2007), outlined several reasons for the early closure of mentoring relationships. These include the mentor's perception of a lack of motivation on the part of the mentee, unfulfilled expectations, a lack of mentor relational skills (for example, inability to bridge cultural or socioeconomic differences) and interference from the youth's family. Working with higher-risk youth could exacerbate any or all of these challenges.

To explore the types of obstacles mentors faced in their meetings and perhaps shed more light on why some decided to close their matches, we asked them about the extent to which they experienced a range of challenges falling into five broad categories: connecting with youth; interacting with and getting support from the youth's family and getting together with the youth (for example, youth being ready for their meetings); the mentor's personal life; managing their mentee's behavioral issues; and bridging cultural or economic differences with the youth or the youth's family (see Table 3.6 on the next page). We also asked mentors to what extent aspects of their match were "unexpected," surmising that to the extent they were, mentors might find this challenging.

Connecting. Many mentors reported having difficulties connecting with youth, and this set of challenges did not vary strongly depending on youth's risk profiles. More than two fifths of mentors reported challenges around having conversations with their mentee (44 percent)[48] or keeping their mentee engaged (41 percent). These kinds of obstacles can make moving forward in a mentoring relationship very difficult. In keeping with these statistics, trouble connecting was a common theme in mentors' responses to our open-ended survey question: More than 60 mentors noted such challenges.

> *I may be too old for Luis to bond with. I think I relate better with someone younger, or someone older than Luis.… Ideally, he should have been matched with a similar ethnic mentor for him to truly bond and grow from a mentor. While I had a similar upbringing, there is a component missing from our relationship that I cannot identify.*

Several mentors noted that their interests or personalities did not seem to match those of their mentee. A little more than a quarter (28 percent) said that this was challenging for them. Although programs reported paying careful attention to common interests when making matches (see Appendix C), 38 percent of mentors did not believe that they shared similar interests with their mentee.[49] And a notable minority of mentors (17 percent) felt that the program did not do a good job matching them. These differences, unfortunately, left some mentors feeling unprepared:

> *I was surprised to be matched with Cristina after being told so many times prior that I would be matched with someone who had similar interests as myself. While I've enjoyed having Cristina as my match, things have been difficult at times, as she has a very different opinion as to what's fun than I do. I'm a very active person, and she prefers to stay indoors; I enjoy sports, and she prefers computer games.… I feel like the program should be more forthcoming in their match making, so volunteers can be better prepared to handle their given situations.*

Differences in interests or personalities were most commonly reported by mentors who were matched with the highest-risk youth;[50] however, experiencing *challenges* around this dissimilarity was just as frequent across all four risk groups. It is likely that common interests sometimes took a backseat to more pressing concerns, as programs felt a need to prioritize other youth or mentor characteristics to make a good match (for example, location, ethnicity, mentee needs, parental preferences). Additionally, youth may not have shared this information during enrollment, and mentees' families may not always have known—or been forthcoming in sharing—their child's interests (a theme we discuss more later in this chapter).

Yet, this was not typically the case. In fact, well over half of mentors (62 percent) *agreed* that their mentee shared their interests, and the vast majority (83 percent) felt the program did a good job matching them with their mentee.[51] Several specifically noted how well matched they were:

> *I am thrilled at how fantastic of a match was accomplished. From day one, Keisha and I have been like two peas in a pod.*

Table 3.6
Mentor-Reported Expectations and Relationship Challenges

Expectations and Challenges	Overall	Association with Risk[a]
Mentor expectations were *not* correct about[b]...		
The mentee's needs	24%	High/High > High Env
The mentee's family's needs	39%	High Env > Low/Low
The time commitment required to develop a strong relationship with the mentee	24%	Low/Low > High Env
Mentor-reported challenges[c]		
Connecting		
Having conversations with my mentee	44%	Low/Low & High/High > High Env
Keeping my mentee engaged	41%	None
Differences in our interests or personalities	28%	None
Youth's family and meetings		
Getting together (for example, transportation, scheduling, disconnected phones)	48%	None
Mentee's preparation for meetings (for example, being ready on time, canceling meetings)	27%	High/High & High Env > Low/Low
Getting support from mentee's family	27%	High/High & High Env > Low/Low
Mentee's family asking for too much help	12%	High/High & High Env > Low/Low
Mentor's schedule or personal situation	66%	None
Managing mentee's behavioral issues	24%	High/High & High Ind > Low/Low
Bridging differences		
Bridging racial/cultural differences	15%	None
Bridging economic differences	21%	High/High > Low/Low

[a] Tests of associations with risk controlled for youth gender, age and race/ethnicity as well as which mentoring program served the match.

[b] Mentor "disagreed" or "strongly disagreed" that his or her expectations were met in the areas listed.

[c] Mentor rated the area as "somewhat" or "very" challenging.

These findings are important, given the recent meta-analysis by DuBois et al. (2011) suggesting that programs matching on similar interests yield stronger effects than those that do not. Such compatibility may very well help matches to flourish and ultimately benefit youth.

Youth's family and meetings. Challenges around meetings and issues with the mentee's family were also fairly common. For example, more than a quarter of mentors (27 percent) noted difficulty getting support from the mentee's family, and 12 percent cited challenges around the mentee's family asking for too much help.

In their responses to our open-ended question, a handful of mentors discussed parents behaving in ways that made it difficult for the relationship to flourish (for instance, not allowing the match to get together or being too involved in the match). Yet, in most cases, mentors simply noted difficulties getting in touch with the family or a general lack of communication between them. More specifically, almost half (48 percent) reported challenges in simply setting up meetings and getting together, and about a quarter (27 percent) said they struggled with their mentee being unprepared for their meetings, not being ready on time or frequently canceling meetings.

It is difficult to ascertain whether these issues reflect a lack of interest on the part of the youth or the youth's family, or whether they simply reflect day-to-day struggles the families were going through—with frequent moves, changing phone numbers, busy schedules, etc. It is noteworthy, in this context, that three of the four challenges under the "family and meetings" umbrella were more common in matches with both groups of youth facing high environmental risk.[52] This suggests that mentoring relationships may be harder to build and sustain when youth's families are under stress from poverty, divorce and other challenges.

The mentor's personal life. Almost two thirds of mentors (66 percent) reported that they faced obstacles in their personal life or in having enough time to devote to their mentoring relationship. Responses to the open-ended question suggested that sometimes this was simply a function of the mentor's other commitments (for example, a job that got "busier"); some mentors also reported that the match required more time than they had bargained for. In some instances, these issues coincided with a match that was already struggling, making it all the more likely that the relationship would end prematurely:

> *I experienced many difficulties in my personal life during the end of the match, loss of income, need to move, work stresses, more than I have ever experienced in my life, so my mentee not responding to me or taking time for me made it easy to let this go.…*

As we discussed earlier, mentors reported most often that they were the instigator of their match's closure. In many cases, matches were closed at least in part because the commitment simply did not fit into the mentor's schedule—a challenge that should be largely avoidable through detailed discussions with the volunteer *before* the match is made about his or her schedule (and how it might change over time).

Managing mentee's behavioral issues. About a quarter of mentors (24 percent) reported that they found managing the mentee's behavior at least somewhat challenging. This coincides with our description of youth characteristics suggesting that sizable minorities of youth were experiencing some sort of individual-level challenges, including 48 percent struggling with a mental health issue and 23 percent engaging in problem behavior. And, in fact, mentors matched with youth who were relatively high on individual risk were most likely to report this challenge.[53]

In some cases, mentors discussed being unprepared to deal with the severe needs of their mentee which included Attention Deficit Disorder (ADD), mental health or behavioral issues and overwhelming family issues (for example, a parent in rehabilitation, children in and out of foster care, abuse). In some cases, these needs were more than mentors felt equipped to handle. As one noted:

> *I had a lot of issues with my match. She wanted to live with me.… She would call upwards of 20 times a day, even when my [case manager], her [guardian] and myself asked her not to.… I think my particular match had a lot of issues that needed to be dealt with and resolved, and no one who I [spoke with] addressed my concerns to give me any support or ideas as to how it should be handled. Eventually I just stopped calling and stopped taking her calls.*

As we discuss later in the concluding chapter of this report, the preceding types of challenges suggest that programs should work up front with mentors to help them anticipate and understand how to deal with youth (and family) needs that can be overwhelming—particularly when challenging behaviors or needs are presented to staff at enrollment. By warning mentors that unexpected challenges will

often arise and teaching them how to help families access appropriate resources, programs may increase the odds that youth are connected with the right kinds of support.

Bridging differences. Most mentors were matched with youth who did not share their ethnic background, and, as discussed in Appendix C, mentors had had very little personal experience with many of the issues their mentees were experiencing. For example, only 12 percent reported having experienced poverty, and two fifths reported that they had not faced *any* of the challenges we asked about in our survey (including family struggles, school challenges and problems with parent or peer relationships). And although 40 percent reported past professional experience with youth (for example, as a teacher or youth worker), and a substantial number had experience with youth facing a range of risk factors, almost a third had *no* experience with youth facing any of the risk factors we outlined. Thus, it would not be surprising if mentors had difficulties bridging differences between their own personal experiences and those of the youth with whom they worked.

Yet, relatively few reported having difficulty bridging racial (15 percent) or economic (21 percent) differences with the youth or the youth's family, and only a couple of mentors noted these types of challenges in their responses to our open-ended question. While challenges with racial/cultural differences generally were reported by similar numbers of mentors regardless of the risks faced by their mentee, challenges bridging economic differences were more prevalent among mentors matched with the highest-risk youth.[54]

Mentor expectations. Reflecting these challenges, a sizable minority of mentors reported that they simply had not expected the kinds of issues they were confronted with when they were matched. Almost a quarter (24 percent) disagreed or strongly disagreed that their expectations prior to being matched were correct about their mentee's needs—a sentiment that was particularly common in mentors meeting with the highest-risk youth.[55] (See Table 3.6.) Almost two fifths (39 percent) had unmet expectations about their mentee's family, a feeling more commonly reported by mentors matched with youth who were high only on environmental

risk.[56] Finally, almost a quarter (24 percent) felt their expectations were incorrect about the time that was required to develop a strong relationship with their mentee.[57] Several mentors, in their responses to our open-ended question, noted that they would have approached the match quite differently and, in their view, more effectively had they known about their mentee's needs prior to being matched.

These findings prompt the question of why mentors were not better informed about the needs of the youth they were matched with. Although parents clearly reported a number of challenges in their children's lives, it is possible that these reports were not always complete. Program staff told us that parents did not always share the extent of their child's difficulties before the child was matched. Perhaps this was in part to ensure that the child was served by the program (parents were not told that programs actually wanted to involve youth experiencing more challenges). In some cases, it may also have been that parents were not fully aware of their children's difficulties or did not understand how such factors could present challenges for a mentor.

One mentor noted the difficulty he faced when realizing his mentee had much more severe needs than he had bargained for:

> *When I was originally matched with John, the program staff had no idea he was currently in and out of the Juvenile Detention Center.… The very first time John and I hung out, outside of our initial match meeting, was me taking him to court and speaking on his behalf to the judge.… So after the first court appearance, I knew this was not going to be a normal 'let's have fun together' match.…*

Clearly, in some cases, mentors and staff did not fully understand the extent of youth's challenges. However, most of the time, mentors reported that staff shared key information with them. Almost three quarters (74 percent) agreed or strongly agreed that staff shared important information with them about their mentee; 19 percent were neutral; and only 7 percent disagreed or strongly disagreed.[58] Taken together, these findings suggest that mentors felt staff were fairly forthcoming with them about their mentee's needs, yet even more information would have been helpful.

It is important to note that although mentors faced significant challenges, a little over half (53 percent) of youth's first matches stayed together for at least a year, and that mentors and youth, on average, reported fairly strong relationships. Thus, in many cases, obstacles were overcome or at least "worked around" in a way that did not impede the development of fairly strong relationships for the average youth.

Summary

On average, mentors and youth in this study appear to have developed fairly strong relationships. Reports of their feelings about each other, youth's reports on how youth centered the relationship was and the extent to which the relationship focused on youth's personal growth or goals all suggest the development of at least moderately high-quality mentoring relationships. Similarly, the types of activities matches engaged in (focused mostly on having fun and getting to know each other) and the goals reported by mentors (most often, to be a friend to youth) are both in line with past research about the conditions that can help matches thrive.

Yet, almost half of the youth had experienced at least one match closure by the time of our follow-up survey (13 months after they entered the program), and only about 60 percent were still in an active match by that time. And although (importantly) match length and quality did *not* differ based on the types of risks the youth were facing when they started the match, there were more subtle differences in the day-to-day interactions of matches with higher- and lower-risk youth that warrant consideration for practice.

Both the challenges mentors reported facing and the reasons their matches ultimately closed varied depending on the youth's risk profile.

For example, mentors reported that they experienced significant challenges in connecting with and getting support from the mentee's family when matched with youth who were relatively high on environmental risk factors (they also reported that these youth cancelled more meetings). This is not surprising—environmental risk captures difficulties in the youth's families or home environments—but it does suggest that preparing mentors *and families* for the relationship up front could go a long way in ensuring that matches succeed.

It is also worth noting that, from the mentors' perspective, one third of matches closed because of a lack of youth interest, and 17 percent closed because the youth did not seem to need a mentor—both reasons were especially likely to be cited by mentors matched with the lowest-risk youth and were cited far more often than youth's needs being too severe. This finding reminds us that volunteers want to feel needed and believe that their time is being used wisely.

Thus, even though our measures were not able to detect differences in relationship quality or length across these risk groupings, mentors' experiences in these matches appear to be quite different based on youth's risk.

The next chapter explores the efforts programs made to prepare the mentors and support these matches and how these efforts were associated with match length and strength.

How Were the Matches Supported?

Chapter IV

Programs in the initiative varied in the types and amount of support they offered to participating matches. And even within each program, matches varied in their experience of key supports—for example, how much ongoing training they took advantage of and the extent to which they felt training and support were sufficient to meet their needs.

Research suggests that these variations may have important implications for the strength and length of mentoring relationships. For example, the amount of pre-match training mentors received has been linked with their reports of relationship quality and, in school-based matches, their continuation of the relationship into a second school year (Herrera et al. 2000, 2007). In one study, mentors' perceptions of training quality were linked with how confident they felt about taking on the mentor role (that is, "mentor efficacy"), which in turn predicted youth reports of relationship quality (Parra et al. 2002).

Studies have found similar links between support—usually in the form of calls from program staff to check in and see how the relationship is progressing—and the strength and length of mentoring matches. Specifically, mentors who said they'd experienced higher-quality or more frequent support from program staff reported stronger relationships with their mentee (Herrera et al. 2000) and were more likely to continue their match (Herrera et al. 2007, 2008). In one study of the BBBSA school-based mentoring program, high school student volunteers were generally found to be much less effective mentors than adults. However, in programs where volunteers reported relatively frequent communication with program staff, high school mentors yielded larger benefits for their mentees in several areas, compared with high school mentors in low-communication programs (Herrera et al. 2008).

This research base informed the design of program enhancements that were integrated into the initiative (see Program Enhancements text box). The two largest participating programs were asked to implement specific programmatic enhancements to increase the support available to matches during

---

Program Enhancements

Two programs in the initiative were asked to implement enhancements to their model for those matches enrolled during the first year of the initiative, when the random assignment portion of the evaluation was taking place at these programs. The enhancements, which were mainly aimed at providing stronger support for matches, consisted of:

1. **Early-match training**[a]
   - Two-hour minimum, group, in-person, curriculum-based training
   - Before or within one month of the match's start date
   - Must include active/experiential learning and practice

2. **Ongoing match support**
   - Monthly contact with mentor, youth and parent/guardian via phone or in person[b]

3. **Match meeting frequency**
   - Matches asked to meet three or more times a month, with each contact lasting at least two hours
   - When a match cannot meet in person, mentors may communicate with youth by phone or (reciprocated) email[c]

4. **Match length**
   - Mentors and youth/parents are asked to commit to 18 months

[a] The early-match training enhancement described here differs slightly from the definition of pre-match training described in the text. The proportion of programs implementing pre-match training and of mentors receiving it, are based on the definition in the text.
[b] Programs were allowed to substitute reciprocated email contact with mentors or parents for phone or in-person communication, but not in consecutive months.
[c] Phone/email communication could not substitute for in-person meetings two or more weeks in a row.

---

the first year of each youth's involvement in the study; and one additional program decided independently to implement these enhancements. It was hoped that these enhancements would help strengthen matches, sustain them over longer periods of time and ultimately yield strong benefits for youth (Grossman, Rhodes 2002).

Thus, in this study, we set out to assess whether various program practices, including the specific enhancements, were linked with stronger and longer mentoring relationships. We also wanted to explore whether these practices were more or less effective for youth with different risk profiles.

In Chapter 3, we reported that mentors working with higher-risk youth experienced distinct challenges—which might require strong training and additional supervision and support to overcome. For this reason, it would not be surprising if certain practices turned out to be particularly important for matches involving higher-risk youth.

In this chapter we examine these issues by using program-reported data for each match to outline the practices implemented by participating programs and survey data from study participants to describe how these practices were experienced by youth, mentors and parents. In doing so, we explore how youth's risk status was associated with the practices (that is, whether matches with higher-risk youth experienced different types and levels of support). We also note associations between program practices and match outcomes including the strength and length of the relationships. In addition, we explore how links between program practices and match outcomes vary depending on youth risk (that is, whether certain practices are more or less important for match outcomes based on youth's risk profile). Finally, we examine whether the abilities of the staff who supervise the matches are linked with mentors' experiences of program support as well as match length and strength.[59]

Training and Support

Programs did a variety of things to prepare volunteers for taking on the mentor role and to support program participants—mentors, youth and their parents—over the course of their relationship.

Mentor training. At the beginning of the study, all but two[60] programs reported that they required pre-match training for their mentors that met the following criteria: (1) It was implemented before the match was made or within one month of the match's meeting; (2) it was either face-to-face or online[61] (that is, not simply written materials or by phone); (3) it was interactive or experiential (for example, involving role play); (4) it was focused on skill-building; and (5) it included various topics of relevance to developing strong, successful matches. The length of training for all but one of the programs was between one and three hours.

The university-based program held a 16-hour training, which included all-day activities in addition to a sleepover. Four programs offered, but did not require participation in, ongoing training for mentors—ranging from 4 to 15 offerings per year. All but one of the programs also had pre-match orientation for youth, with two implementing more intensive youth training (similar to that described above for the mentors). One of these two programs conducted "Big-Little trainings" in which mentors and mentees attended a training jointly.

By the end of the initiative, the programs had expanded their pre-match mentor training in several ways. By this time, all reported having pre-match group-based mentor training using most of the criteria noted above. This includes two programs that initially were resistant to implementing pre-match (or "early-match") in-person training because they believed they would have difficulty getting mentors to enroll if they required it. Both now require it (albeit not within four weeks of the beginning of the match). All but one of the programs also lengthened its pre-match training so that, by the end of the initiative, training across the programs ranged from two to seven hours. Three had begun using external trainers in addition to internal staff. And five of the seven programs had implemented stricter requirements around attendance. Finally, by the end of the initiative, all but one program offered ongoing training to their volunteers.

The programs in the study entered detailed information about each match into a database on a monthly basis. These data showed that 56 percent of all mentors received early-match training that met all the criteria outlined above—other than the requirement that training occur within one month of the start of the match; when we consider only training that occurred within the first month, the percentage drops to less than half (46 percent).[62] On average, early-match training for the mentors matched in our study lasted between two and three hours.[63] Mentors of youth high only on individual risk were significantly more likely to receive early-match training than mentors of youth in the other three risk groups.[64] This difference could be the result of programs making a special effort to ensure that the mentors of youth with personal challenges

received training. The mentors of these youth may have also taken more initiative to seek out this type of support. Among mentors who received early-match training, there were no differences in the number of hours of training received depending on risk group, reflecting the fact that training was offered in a group context and thus programs could not tailor it to the needs or preferences of particular mentors.

Again based on program-reported data for each match, a little less than a third (31 percent) of mentors received ongoing training (that is, training that was intended to assist mentors after their matches had begun).[65] Among those who received any ongoing training, the amount received was about four hours on average. Mentors of youth who were relatively high only on environmental risk were more likely to receive ongoing training than mentors of the lowest-risk youth.[66] Also, among those who received it, mentors of the highest-risk youth (that is, those high on both individual and environmental risk) received slightly more hours of training.[67]

Overall, 64 percent of mentors received some form of formal training, either before or after their match was made; thus, approximately one third of mentors received no formal training.[68] Mentors of youth who were high only on individual risk were significantly more likely to have received some training than mentors of youth in the lowest-risk group.[69]

Volunteer training and support covered a range of topics, as shown in Table 4.1, and mentors reported that training on some of these topics was more essential to their work with youth than others. Mentors' most common self-reported training/support need was related to strengthening their mentoring relationship. The second most common need, reported by 58 percent of mentors, was learning how to interact with the youth's family. Mentors matched with the highest-risk youth were most likely to report this training need.[70] The fact that more than half of all mentors reported needing help in this area echoes findings presented in Chapter 3 about the challenges mentors faced in their work with youth and why matches ultimately

Table 4.1
Training/Support Topics

Training/Support Topics	Percent Needing Help in This Area	Percent of Those Who Needed Training/ Support Who Said It Was Received and Helpful	Differences in Needing Help Based on Youth's Risk Profile[a]
Strengthening the mentoring relationship	67%	79%	None
Interacting with youth's family	58%	72%	High/High > Low/Low
Addressing youth's emotional issues	51%	69%	High Ind > Low/Low; High/High > Low Ind
Addressing youth's social issues	50%	70%	High Ind > Low/Low; High/High > Low Ind
Addressing youth's behavioral issues	47%	69%	High/High > Low Ind
Interacting with youth of different ages	43%	71%	None
Working with higher-risk youth	42%	70%	High/High > Low/Low
Increasing comfort with youth's socioeconomic status and culture	38%	74%	None
Navigating social service systems (for example, welfare agencies, juvenile justice, foster care)	24%	57%	High/High > High Env

Note: See Designations of the Four Risk Profiles text box in Chapter 3 for definitions of each of the risk categories noted in the last column of this table.

[a] Tests of associations with risk controlled for youth gender, age and race/ethnicity as well as which mentoring program served the match.

ended: Navigating the youth's family and ensuring their support appears to have been difficult for these mentors and, for a substantial minority, a threat to sustaining the relationship. This training need and its related challenges highlight the importance of preparing *both* the mentors and the families they will be working with. Only one of the seven programs had pre-match training for parents when the study began; four offered an orientation; and two offered neither. By the end of the study, one of the programs had added parent orientation to its practices.

Sizable groups of mentors reported needing help addressing youth's behavioral (47 percent), social (50 percent) or emotional (51 percent) issues. Not surprisingly, mentors matched with youth who were high on individual risk were particularly likely to report wanting help in each of these three areas.[71,72]

About two fifths (42 percent) of mentors reported needing training or support around working with higher-risk youth. As might be expected, mentors matched with the highest-risk youth were particularly likely to report needing this help.[73] About a quarter of mentors (24 percent) reported needing help navigating social service systems—again, a need that was particularly common in mentors matched with the highest-risk youth.[74]

These differences in mentor-reported training needs suggest that mentors could benefit from training and support that is tailored to the specific type of youth with whom they are matched. As we discuss in more depth in Chapter 6, these characteristics could be gleaned by programs before the match is made and then used to design more targeted trainings.

For most of the topics we asked about, more than two thirds of mentors felt that the training/support they had received was helpful. And, importantly, very few mentors across all nine topics (8 percent or less for each topic) reported that they had needed help in a given area but that it was *not* available.[75] At the same time, for the 30 percent or so of mentors who reported having a need for training that was *not* addressed (that is, training was needed but the mentor reported it was either not available or that it was provided but not helpful), they likely felt

somewhat unprepared to deal with the challenges their mentee presented. This may have been particularly true for mentors matched with the highest-risk youth, as they were the most likely to report needing help in several areas.

Mentor support. Before the initiative began, all of the participating programs had expectations for how often staff should contact mentors. Six of the seven programs reported that they contacted mentors once a month during the first year of the match; one reported communication every two months.[76] All but one program also required regular communication with parents and youth, but less frequently (only one required monthly communication; the others required communication every two months).

Based on the data collected by the programs for each match,[77] about three fifths (61 percent) of mentors communicated with program staff during at least 70 percent of the months that their mentoring relationships were active. Interestingly, a notably larger proportion (89 percent) of those who completed the mentor survey agreed that they were contacted at least monthly by program staff. Whereas mentor reports may have reflected the efforts of staff to reach them (for example, leaving messages, sending emails), even when unsuccessful, program reports were limited to instances of actual contact. It is thus possible that mentor reports reflect more frequent efforts of staff to *try* to contact them on a regular (monthly) basis. This possibility highlights the role mentors (and other program participants) can play in determining whether they actually receive support from staff. That is, because mentors determine whether support calls are actually answered or returned, the frequency of support actually *received* is by no means driven only by the program. This is important to keep in mind when considering other findings in this chapter.

Mentors reported that these contacts (which typically occurred over the phone) were not very long: 27 percent said they were 5 minutes or shorter. Another 47 percent said they generally were 6 to 10 minutes. Only about a quarter (26 percent) reported that they generally lasted longer than 10 minutes. The length of these support calls

differed by the youth's risk status. Mentors matched with youth who were high on only individual risk reported longer calls than those matched with youth low on individual risk.[78] These mentors may have sought out more in-depth help to address their mentee's social, emotional or behavioral needs, or their case managers may have known that they needed more help and thus spent more time with them.

Despite these support contacts being fairly brief, most mentors felt they were helpful in strengthening their match, with 30 percent reporting they were very helpful and about half (54 percent) reporting they were somewhat helpful. Only 16 percent felt they were "not very" or "not at all" helpful. The overwhelming majority of mentors also agreed or strongly agreed with each of the following statements about the support they received from program staff: "program staff seem willing to help me" (90 percent); "program staff have shared important information about my mentee" (75 percent); "program staff have given suggestions on what I can do with my mentee" (88 percent); and "program staff seem truly concerned about how well our match is going" (85 percent).

In addition to regular communication with program staff, three of the largest programs offered mentor support groups for their volunteers, holding anywhere from 4 to 12 meetings throughout the year. But programs reported that only a small proportion of their mentors attended (less than 15 percent on average). In our mentor survey, 22 percent reported that they had taken part in a mentor support group; 34 percent said the program did not provide such a group; and 44 percent simply did not participate in those that the program provided. Almost all mentors (94 percent) who did participate found them at least somewhat helpful.

This last finding suggests that programs should consider making such groups a more regular part of the support that they provide.[79] Programs noted informally that getting mentors to participate in groups like this once their matches have started is challenging. However, if attending at least one group is presented as a requirement, perhaps mentors would attend one or more, see their value and continue participating.

Mentors paired with youth experiencing relatively high levels of individual risk were more likely to receive regular support contacts from programs (with "regular" defined as 70 percent or more of their months with their mentee).[80] As noted, because the amount of support received is driven both by the program (that is, how often the case manager tried to contact the mentor) and the mentor (that is, how often the mentor returned those calls), this difference could reflect both staff's understanding that these mentors needed more help and the mentors simply being more receptive to these calls given their more pressing needs. We found no notable differences in mentor reports of their experiences of the helpfulness of support based on youth risk.

Finally, we also asked mentors whose match had closed about how program staff may have helped them manage their closure. Mentors did not report receiving extensive support in this area. About a third (31 percent) met with program staff to facilitate their match closure. A little less than half (48 percent) reported that program staff offered ideas to help them deal with the match closure, and about 41 percent reported that program staff gave them ideas to help *their mentee* deal with the match closure. Yet, three quarters felt that staff were at least somewhat helpful in making their match closure a positive experience for them.

Youth support. Based on program records, only about a third (34 percent) of youth received regular support contacts from program staff. Interestingly, however, more than three quarters (77 percent) of youth reported that they talked with someone at the mentoring program regularly about how things were going with their mentor, perhaps reflecting their experience of more informal contacts with staff in the context of seeing them at program-sponsored activities and events or simply a more liberal definition of "regular" support contacts than the one we used. Of further note, 84 percent of youth reported that there were staff at the program they could go to if they had problems with their mentor. Interestingly, youth's responses suggested that, in many cases, they had developed a "close" relationship with the case manager. For example, 59 percent said they felt "close" or "very close" to someone who works at the mentoring program; 83 percent said they felt comfortable talking with program

staff; and 71 percent said they could talk with program staff if they had problems at home or school. Whether youth received regular support contacts based on program records did not vary in relation to their risk status. However, the rates at which *youth* reported talking regularly with program staff about their mentor and the percentage who felt they could talk with program staff about problems at home or school did vary depending on youth's risk status. Youth who were high on only environmental risk reported higher levels of support on both measures than youth in other groups.[81] (The other three youth-reported support variables did not differ by risk status.)

Parent support. We were also interested in the types of supports available to the youth's families. Based on program records, about half (48 percent) of parents received regular support contacts from program staff. Whether parents received this level of support did not differ by youth's risk status. For parents from the two largest programs, we were able to ask them directly about their experiences of support.[82] Nearly all (95 percent) reported that there was someone they could go to at the program if they had concerns about their child's mentor; and most (87 percent) agreed that program staff had been good at listening to their suggestions or concerns.[83]

Were Program Practices and Experiences Associated with Stronger, Longer Matches?

The programs were selected for involvement in the initiative in part because they were already implementing research-based practices and thus had some evidence that they could potentially serve higher-risk youth in a way that would make a difference in their lives. In spite of this baseline of strong practices, there was still variability across programs in how they trained and supported their matches. The enhancements implemented at three of the programs strengthened some of the supports they offered beyond what they typically provided to youth. In addition, some programs changed practices over the course of the initiative, such as the increased emphasis on early-match training noted earlier in this chapter. These and other sources of variability (for example, changes in program staff) allowed us to examine whether

certain practices, as experienced by individual matches, were associated with stronger outcomes for those matches involved.[84]

Table 4.2 on page 47 summarizes the findings reported earlier in this chapter about the extent to which programs implemented different practices, as well as the frequency with which programs offered matches the full set of programmatic enhancements (that is, early-match training at any time, and match support provided to mentor, youth and parent, 70 percent or more of their months together).[85] As can be seen, this combination of practices was implemented with only 16 percent of matches. When considering matches that received at least early-match training and regular mentor support, but not necessarily regular support contacts for youth and parents (that is, a "partial" set of enhancements), the percentage receiving enhancement practices more than doubles, reaching 35 percent. Whether matches received the full set of enhancements did not differ according to risk status. However, when using the more liberal criteria, we found that matches with youth high on only individual risk were more likely to experience the enhancements than those with youth in each of the other risk profile groups.[86]

A key question in the initiative was whether program practices were related to better match outcomes, including match strength and length. However, analyses addressing these questions are more preliminary than other analyses presented thus far and should not be construed as testing a *causal* association between practices and outcomes. As noted, some mentors take advantage of program practices (for example, participating in ongoing training, returning support calls) more than other mentors. And the same characteristics that lead them to get more of these supports (for instance, strong attachment to their mentee, dedication to the program, persistence) could very well help them to create stronger, longer matches. Thus, the mentor's characteristics—not necessarily the program supports themselves—may be what causes a positive outcome. Similarly, a struggling match could lead a mentor to seek out more support than would be sought by mentors in more successful matches; again, the support does not *cause* the match to struggle, but these matches may end up with higher levels of support. Despite these

Table 4.2

Associations Between Program Practices and Match Length and Strength

Program Practice	Overall Percent Receiving Practice	Links Between Receiving Practice and Risk?	Links with Meeting Frequency?[a]	Links with Match Lasting 12 Months or Longer?[b]	Links with Relationship Quality?[c]
Received early-match training (ever)	56%	Yes	Yes	Yes	Yes: YC, GR, CL
Attended ongoing training	31%	No	No	Yes	Yes: GR, CL
Mentor received match support 70% of months or more	61%	Yes	Yes	Yes	No
Youth received match support 70% of months or more	34%	No	Yes	No	No
Parent received match support 70% of months or more	48%	No	Yes	No	No
Received partial enhancement package[d]	35%	Yes	Yes	Yes	Yes: YC, GR, CL
Received full enhancement package[e]	16%	No	Yes	No	Yes: CL

Note: These analyses controlled statistically for the youth's gender, age, race/ethnicity, risk status group and program and, in the case of analyses predicting relationship quality, mode of survey administration at follow-up (in-person or phone). Unless otherwise noted, findings refer to practices provided to the match that was currently or most recently active for the youth at the time of the follow-up assessment.

[a] We designated "frequent" meetings as three or more times a month at least 70 percent of the match's months together.

[b] Analyses for this match outcome were limited to the youth's first match because not all of youth's subsequent matches were followed long enough to be able to assess whether they lasted 12 months.

[c] We tested associations with the following youth-reported measures of mentoring relationship quality: growth/goal focus (GR), youth centeredness (YC) and closeness (CL). All associations were in the expected direction (that is, more support/training was linked with higher scores on the measures).

[d] For a match to be designated as receiving the partial enhancement package, the mentor needed to receive early-match training (at any time) and regular support (70% of months or more), but there did not need to be regular support provided to the youth and parent.

[e] For a match to be designated as receiving the full enhancement package, the mentor needed to receive early-match training (at any time) and the mentor, parent and youth needed to receive regular support (70% of months or more).

inherent limitations, we believe our preliminary analyses are useful, suggesting which practices may have the most potential to strengthen matches.

Table 4.2 displays how key match outcomes (that is, how frequently the match met, match length and relationship quality) varied with these practices. We found associations between program practices and all three of these important match outcomes. First, mentors who met with their mentee regularly (that is, at least three times a month during 70 percent or more of their months together) were more likely to: (1) have received early-match training;[87] (2) have received regular support;[88] (3) be matched with youth who received regular support[89] and whose parents received regular support;[90] and (4) have received either the partial or full set of program enhancements.[91] These findings suggest that program training and support may encourage matches to meet more regularly. However, as noted, program training and support are not completely program-driven practices. For example, mentors and families must return calls in order for support to be delivered—something programs noted was often very difficult, particularly for families under stress. Thus, the same obstacles that may prevent families and mentors from returning calls (for example, a lack of interest, higher stress

levels) may also prevent matches from meeting more frequently. In spite of this caveat, the findings do suggest that this is an important area for others to examine more rigorously.

Second, matches that lasted 12 months or longer over the course of the initiative (that is, not just by the 13-month follow-up)[92] were more likely than shorter matches to have received almost all of the mentor-directed supports that were tracked by programs. The mentors in these longer-lasting matches were more likely to have received early-match[93] and ongoing[94] training as well as regular support.[95] And these matches were more likely to have received the partial set of program enhancements.[96] Again, these links likely reflect both the fact that more support and training can encourage matches to stay together and that matches that are on the road to staying together are more likely to invest in some of these supports.

As can be seen in Table 4.2, we also found associations between youth-reported relationship quality and three of the program practices we tracked. Youth whose mentor received early-match training reported higher levels of closeness, a more youth-centered approach by their mentor and a stronger growth/goal orientation to their match (that is, all three of the youth-reported relationship measures examined). Similarly, youth whose mentor received ongoing training reported greater feelings of closeness toward their mentor as well as higher levels of growth or goal orientation to their matches. In a nutshell, having a mentor who received early-match or ongoing training was predictive of youth feeling more positive about the relationship at the 13-month follow-up.

In addition, youth whose match received the partial enhancement package reported feeling significantly closer to their mentor as well as more youth centeredness and growth/goal orientation in their match than did youth whose match did not receive these supports. Youth who were in a match that received the full enhancement package reported greater feelings of closeness toward their mentor.

Finally, we also tested whether practices were associated with the match outcomes in Table 4.2 *differentially* based on the risk status of the youth. For example, is support or training key for match

outcomes but only for the highest-risk youth? We found no evidence to indicate that this was the case. The associations reported in Table 4.2 between program supports and match outcomes were relatively consistent across the four risk groups.

Case manager characteristics. The case manager is the main point of contact between the program and all individuals related to the match—the mentor, youth and parent—potentially making this staff person one of the most important influences the program has on the match's progress. The approach the case manager takes and his or her background, skills and persistence in reaching matches could be important influences on match success. Practitioner experience supports this notion, but very little research has actually tested it.

To understand more about how case manager characteristics might affect match success, we administered surveys to each case manager's supervisor asking about the case manager's background, caseload and approach to working with youth. We then examined whether case manager competence (as reported by supervisors) was linked with mentors' experiences of support, including the extent to which the case manager seemed truly concerned about how well the match was going, had shared important information about the youth and had offered suggestions for activities. We found that, indeed, supervisors' reports of case manager competence were associated with mentors' reports of the quality of the support they received. Mentor reports of greater support quality were, in turn, associated with youth reports of better relationship quality (feelings of closeness, youth centeredness and growth/goal orientation) and with the length of the match at follow-up. These findings suggest that the associations between case manager competence and match outcomes may result, at least in part, from how the mentor experiences the support provided by this staff member. There were no indications that the links we measured varied based on youth risk.

Overall, these findings support the idea that more capable case managers are able to yield matches that are higher in quality (as experienced by youth) and that last longer. Moreover, they may do so at least in part by supporting mentors in a way that *mentors* feel is helpful—specifically that the mentor

feels the case manager is concerned about the success of the match, helps the mentor learn more about the youth and offers suggestions for match activities. (See Appendix G for more details.)

Summary

Our analyses revealed that matches received fairly similar types and amounts of program support regardless of youth's risk status, with one important exception. Mentors paired with youth who were high on individual risk were more likely to have received several program supports. For example, they were more likely to have had early-match training and more consistent support contacts from their program; they also reported longer support calls. It is hard to know if these findings were the result of mentors looking for more help, programs offering more help or some combination of the two.

Mentors' self-reported training/support *needs* differed to a much greater extent depending on their mentee's risk profile. For example, mentors paired with the highest-risk youth were more likely to say they needed help learning how to interact with the youth's family. Mentors who were matched with youth high on individual risk reported greater concerns about youth's behavioral, social and emotional issues. Together, these findings suggest

mentors might benefit from training that is tailored to the specific youth with whom they're matched.

Finally, our study found associations between program practices and important match outcomes. Case managers' competence was linked with mentors' experience of program support, which was, in turn, linked with youth-reported relationship quality and match length. Training and regular support contacts (for mentors, youth and parents) were associated with various indicators of match success. And the combined presence of two of these practices—early-match training and consistent support contacts for mentors—was linked with matches meeting more regularly, lasting at least 12 months and being rated as stronger on multiple dimensions by youth. These findings should be interpreted cautiously, as the qualities that lead mentors to seek out more support may also result in stronger matches (that is, might not be the result of the additional support). Still, we believe these findings provide good preliminary evidence about program practices that may effectively nurture mentoring relationships and ultimately benefit youth.

In the next chapter, we examine the effects of program involvement on participating youth, and how these outcomes may have differed for youth with different risk profiles.

How Did Youth Benefit?

Chapter V

As discussed in Chapter 2, the mentoring programs in the initiative were able to enroll and serve youth experiencing a wide range of risk factors. A question of central importance, of course, is whether the programs made a positive difference for these youth. Prior studies have indicated that volunteer-centered, one-to-one community-based mentoring can impact an array of youth outcomes. In this evaluation, we considered the effect of mentoring in eight broad areas:

- Emotional/psychological well-being;

- Adult relationships;

- Peer relationships;

- Academic attitudes;

- Academic performance;

- Academic behaviors;

- Antisocial behaviors; and

- Prosocial (positive social) behavior and activities.

The first set of analyses reported in this chapter examines outcomes at follow-up using an "intent-to-treat" design[97] for two distinct groups of youth—those who had access to a mentor as part of the random assignment portion of the study, and those who had access to a mentor as part of the quasi-experimental portion of the study (described in Chapter 1)—and how these outcomes differed from those for youth in the control/comparison group. We then turn to the question of whether program effects differed based on our four risk profiles (youth relatively high on both individual and environmental risk; youth relatively low on both types of risk; youth high only on individual risk; and youth high only on environmental risk). Here, to increase our ability to detect differences, we combined all youth with access to a mentor across both the random assignment and quasi-experimental groups.[98]

Chapter 3 showed that the average youth in this study reported having fairly strong relationships with their mentor regardless of risk level and that matches were similar in duration across youth in the four risk profile groups. Because match strength and length have been linked with youth outcomes in prior research (Grossman, Rhodes 2002; Herrera et al. 2007; Rhodes, DuBois 2006), one might expect program effects to be similar for the four groups in our study. Yet, as also shown in Chapter 3, youth's and mentors' experiences in the program appeared to vary in noteworthy ways depending on the risk factors youth were facing. For example, the mentors of youth with relatively high environmental risk tended to report different kinds of challenges than the mentors of youth with relatively high individual risk, and their matches, when closed, were reported to have ended for different reasons. Such differences could have affected the day-to-day tenor of the relationships and their ultimate impact on youth outcomes in ways not captured by youth or mentor reports of relationship quality.

In addition to examining program effects for youth in the four risk profile groups, we also examined how effects may have varied by other youth characteristics and experiences—such as gender, race/ethnicity and whether the youth lived in a single-parent home. Together with the results of the risk profile analyses, the findings from these analyses provide information that can potentially help programs make decisions about whether and how to target their efforts toward specific groups of youth.

All of the analyses in this chapter are based on youth for whom follow-up survey data were available. Rates of success in obtaining follow-up data were fairly consistent across the random assignment treatment group (81 percent), the quasi-experimental participant group (79 percent) and the control/comparison group (83 percent).[99] (More details on the methodology used to test for program effects on youth outcomes are provided in Appendix B.)

Program Effects on Youth Outcomes

Below are the specific measures we examined as primary outcomes of interest, within each of the broad areas noted at the beginning of this chapter (for examples of questions from each scale and information on scale reliability and other secondary outcomes in each of these areas, see Appendix A):[100]

- Depression (that is, depressive symptoms—a measure of emotional/psychological well-being);

- Parent trust (a measure of adult relationships);

- Social acceptance (a measure of peer relationships);

- Self-perceptions of academic abilities (a measure of academic attitudes);

- Grades (a measure of academic performance);

- (Initiation of) skipping school (a measure of academic behaviors);

- Misconduct (a measure of antisocial behaviors); and

- Prosocial behavior (a measure of prosocial behavior and activities).

In addition to these specific outcomes, we created two aggregate measures for each youth: the number of our primary outcomes (excluding skipping school)[101] that showed meaningful (that is, statistically reliable) change in a positive direction over the follow-up period, and the number of key outcomes that showed meaningful deterioration over the follow-up period. We developed these measures based on the premise that mentoring is typically not an intervention with one or two specific targeted goals. Rather, it is a broad-based intervention that is believed to improve youth's lives by addressing the *specific* and differing needs of participating youth across a wide range of areas. Such benefits can be masked when assessing any given outcome because that outcome may present a salient need and area of change for only a subset of youth. Likewise, it is also possible that the broad orientation of mentoring programs may increase the likelihood that youth will exhibit significant improvements (or avoid deterioration) in multiple, potentially distinct outcome areas (for example, depression and grades in school). The aggregate measures allowed us to address these possibilities by testing whether, across the set of primary outcomes, program participants improved in more areas (and experienced setbacks in fewer areas) than comparison youth.

Table 5.1 shows how each of the two mentored groups (that is, youth with access to a mentor) scored relative to the comparison group on each of the primary outcomes at baseline (see Appendix B for information on treatment-, participant- and

Table 5.1
Average Baseline Scores on Primary Outcomes for the Treatment-, Participant- and Comparison-Group Youth in the Analysis Sample

Scale/Indicator	Random Assignment Treatment Group (*n* = 308)	Quasi-Experimental Participant Group (*n* = 615)	Comparison/Control Group (*n* = 321)	Difference Between Random Assignment Treatment Group and Comparison/ Control Group?	Difference Between Quasi-Experimental Participant Group and Comparison/ Control Group?
Depression	1.56	1.45	1.56	No	Yes
Parent trust	3.37	3.43	3.38	No	No
Social acceptance	2.64	2.70	2.62	No	No
Self-perceptions of academic abilities	2.88	2.89	2.85	No	No
Grades	3.55	3.61	3.53	No	No
Skipping school	7%	4%	5%	No	No
Misconduct	12%	11%	11%	No	No
Prosocial behavior	3.56	3.58	3.55	No	No

Note: Tests for significant group differences were conducted in analyses controlling for several key background characteristics of youth (age, gender, race/ethnicity and risk profile), mentoring program to which the youth applied and mode of survey administration (in-person or phone) at baseline. The averages reported for each measure are model-estimated means after accounting for these controls.

comparison-group scores on secondary outcomes and their respective background characteristics at baseline). In all of the outcome analyses presented in this chapter, we controlled statistically for (that is, held constant) any measures that differed between either of the mentored groups and the comparison group at baseline, as well as the mentoring program to which the youth applied, the mode of survey administration at baseline and at follow-up (in-person or telephone) and key background characteristics (see Appendix B for the full list of controls). This helped ensure that we were comparing groups that were as similar as possible—except for whether or not they had access to a mentor.[102]

What Were the Effects of Mentoring 13 Months After Enrollment?

As is true in most community-based mentoring programs, it often took a month or two to find an appropriate mentor for a youth, and many matches closed before 13 months had passed. Thus, the average length of the youth's first match at follow-up was not 13 months but about 8.95 months, as described in Chapter 3.

Table 5.2 on the next page shows the difference in each primary outcome associated with being in the program during the study period for both the randomly assigned youth and the youth who were part of the quasi-experimental study relative to the comparison/control group (estimates of program effects for secondary outcomes are provided in Appendix B). To compare the size of these differences across outcomes and with those reported in prior research, we followed a common research practice of reporting the differences as standardized differences, also commonly referred to as "effect sizes."[103]

At the end of the year (that is, about 13 months after their baseline assessment), compared with similar youth in the comparison group, youth in one or both of the mentored groups reported:

- Fewer depressive symptoms;

- Greater acceptance by their peers;

- More positive beliefs about their ability to succeed in school; and

- Better grades in school.

We found no significant differences in their reports of parent trust or in any of the positive or negative behaviors we assessed—that is, misconduct, prosocial behavior or their likelihood of starting to skip school.

In the two final rows of Table 5.2, we see the differences between each of the two groups of program participants (that is, the treatment and participant groups) and the comparison/control group on the two aggregate measures we created. Outcomes moved positively for significantly more of the key outcomes for both groups relative to youth in the comparison group.[104,105]

To better understand this finding, we also examined the likelihood of youth showing positive change on at least one outcome measure and on multiple outcome measures.[106] Our analyses revealed that, relative to the comparison group, both mentored groups were more likely to improve on at least one measure and to show improvement on multiple outcomes, although in general very few youth in any group exhibited this latter pattern. Significant program effects were not found on the measure that indexed the number of outcomes that deteriorated during the study period.[107]

The largest estimated program effect across all of the outcomes and most consistent across the two mentored groups was in the youth's self-reported symptoms of depression—with an effect size of -.32. As reported in Appendix B, we also found a significant program effect (effect size of -.19) on a parent-reported measure of emotional difficulties that was administered at follow-up only in the random assignment portion of the study (this measure was not included among the primary outcomes reported in this chapter). Together, these findings provide strong evidence that mentoring helped to ameliorate and/or prevent the emergence of emotional difficulties for participating youth.[108]

Very few evaluations of mentoring programs have examined the intervention's effect on depression. One notable exception is Bauldry's (2006) study, which found that program-based mentoring was associated with lower levels of depression for youth who had been involved in the juvenile justice system. Benefits in this area are consistent with what we know about mentoring; it makes sense

Table 5.2
Estimated Program Effects on Primary Outcomes

Scale/Indicator	Random Assignment Treatment Group vs. Comparison/Control Group	Quasi-Experimental Participant Group vs. Comparison/Control Group
Depression	-.14*	-.32***
Parent trust	.08	.09
Social acceptance	.05	.22**
Self-perceptions of academic abilities	.06	.16*
Grades	.07	.19*
Skipping school	-.21	.20
Misconduct	.07	.04
Prosocial behavior	.00	-.02
Number of outcomes showing positive change	.15*	.32***
Number of outcomes showing negative change	-.06	-.12

Note: Standardized mean difference estimates of effect size are reported in the table. These effect estimates were derived from regressions that included the control variables noted in the text (see Appendix B for details).

*** $p < .001$ level of significance.

** $p < .01$ level of significance.

* $p < .05$ level of significance.

† $p < .10$ level of significance.

that youth would simply "feel better" with positive attention from a caring adult and the increased opportunity to engage in fun activities that such a relationship could provide. More targeted outcomes (like improvements in grades or decreases in negative behaviors) likely occur in *some* youth who may need help in these specific areas, but may also take time to emerge. A more emotionally based outcome like depression would appear to be a good bet for a benefit that occurs relatively early on for many youth, and could then serve as a catalyst for further improvements. Although our assessments were not meant to diagnose clinical depression, responses by a substantial portion of our sample (22 percent) suggested they were experiencing worrisome levels of depressive symptoms at baseline,[109] and childhood depression is linked with suicidal behavior, academic and social difficulties, and increased risk for substance abuse and teen pregnancy (for a review, see Malhotra, Das 2007; see also Cash, Bridge 2009). Thus, the improvements we saw in this area are particularly noteworthy.

Youth also experienced improvements in their sense of social acceptance among their peers, although the absolute size of this improvement was small and not statistically significant in the random assignment group. Thus, this study provides more moderate evidence of benefits in this area. Because mentoring is a relationship-based intervention, youth's feelings about other relationships and how they approach them may improve for many participating youth, and (like reduced depression) those changes may serve as the impetus for changes in other domains (Rhodes 2005). Improvements in peer relationships are also significant for youth's later adjustment. Low peer acceptance (or social isolation) is predictive of dropping out, criminal behavior and physical health problems later in life (Parker, Asher 1987; Caspi et al. 2006), again, underscoring the value of impacts in this area.

We also looked at several measures related to academics. The quasi-experimental participant group felt significantly more confident about their

schoolwork at follow-up than members of the comparison group and also reported getting significantly better grades. However, these differences were not significant in the random assignment portion of the sample. Thus, we interpret them with more caution and conclude that the study provides modest evidence of effects in academic areas. P/PV's 1995 study of BBBSA community-based mentoring also found significant positive impacts on youth-reported measures of both of these outcomes (Tierney et al. 1995). Even when mentors only infrequently focus explicitly on schoolwork (as was true for the matches in this study), their interactions with youth may nevertheless be helpful in laying the foundation for youth to have more confidence and success in this important arena. For example, mentors may encourage youth to believe that they can improve their academic abilities with practice, thus fostering the type of "growth mindset" that has been linked with increased success in school, especially among groups that are subject to negative stereotypes regarding their academic potential (Good et al. 2003; Blackwell et al. 2007). The largely indirect nature of such influences on academic outcomes could help account for why the significant effects that were observed on these measures were relatively small, as well as why such effects failed to reach significance in the random assignment portion of the study.[110]

In contrast to the preceding areas, we found no evidence—at least according to youth reports—that mentoring significantly changed youth's behaviors—either by reducing antisocial behaviors or by increasing prosocial ones. We did find a significant difference favoring the treatment group on a measure of *parent*-reported conduct problems (effect size = -.13), a measure we obtained at follow-up only for youth who participated in the random assignment portion of the study (see Appendix B). Yet, because youth and parent reports do not "agree" on this improvement, the findings as a whole do not provide strong evidence that youth benefited in this area.

Our lack of evidence for improvements in youth behavior could reflect the fact that these matches had, on average, just reached the nine-month point (that is, many relationships did not stay together through the entire 13-month period during which we studied them). Grossman and Rhodes (2002) and Grossman and Johnson (1999) found in their

work that impacts grew over time, not appearing until the matches had lasted at least six months. As proposed in Rhodes' (2005) influential model of youth mentoring, a positive connection with a mentor may work first by changing a youth's attitudes toward herself and others, and then, only later, contribute to behavioral changes. In line with this view, a meta-analysis of studies of mentoring for both youth and adults (Eby et al. 2008) found that mentoring was associated more strongly with attitudinal outcomes than it was with behavioral, health or career outcomes. As these researchers noted, attitudes may be more amenable to change than outcomes that are more contextually dependent or more influenced by stable personal characteristics. For instance, a youth's decision to engage in misconduct may be strongly influenced by contextual factors such as peer pressure and parenting, making it more difficult for a mentoring relationship to have a substantial impact on this outcome, at least in the short term (Eby et al. 2008). Stable personal characteristics like gender may also be linked with socialization experiences or broader societal influences that are difficult to offset through mentoring alone. As suggested earlier, mentoring may affect some outcomes across most youth involved (for example, depression), whereas in specific areas that are not of immediate concern for all participating youth, mentoring may tend to benefit only some youth, at least in the short run.

This could also help account for the lack of program effects we saw on the quality of youth's relationships with their parents. Especially in community-based mentoring programs, where participation typically depends on parent referral (which itself can be a clear sign of concern for a child's well-being), there may be less need or room for improvement in this area for many youth. In keeping with this possibility, Table 5.1 shows that average scores on parent trust at the start of the study were not far from the maximum possible score (4.0).

Differences in findings across the two (random assignment and quasi-experimental) samples could reflect a number of issues. First, random assignment (that is, an experimental design) typically yields a more accurate estimate of a program's impacts (Flay et al. 2005). Although we tried to statistically account for any potential differences between the participant and comparison groups

in the quasi-experimental portion of our analyses, these groups may have differed in unmeasured ways that we could not account for. Second, keeping in mind that the experimental portion of the evaluation took place during the earlier portion of the initiative, differences in findings from those in the quasi-experimental portion of the study also could reflect changes over time in the types of youth or mentors who participated in the programs, the economic context in which youth were enrolled or the staff who served them (see endnote 102 for further discussion of these potential changes over time and implications for our findings). It is also possible that improvements in program practices that occurred over the course of the study helped drive differences in findings between the two study components. For example, programs improved their training efforts over time, as described in Chapter 4. Such differences certainly could have contributed to different findings across the two portions of the evaluation.

To summarize, our findings suggest that, to varying degrees, volunteer-centered, one-to-one community-based mentoring had a beneficial effect on youth's emotional/psychological well-being, relationships with peers, confidence and success in school and the number of areas (regardless of the specific outcome) in which youth exhibited meaningful gains. We found the strongest evidence for benefits in depression and in the number of areas in which youth showed improvement.

More moderate evidence was found for benefits in peer relationships, self-perceptions of academic abilities and grades. Using a framework similar to that developed by the well-respected What Works Clearinghouse of the Institute of Education Sciences in the US Department of Education (*WWC Procedures and Standards Handbook* 2011), our findings would indicate "positive effects" on depression and the overall number of positive changes across outcome areas, "potentially positive effects" on perceptions of social acceptance, academic ability and grades, and "no discernible effects" for the remaining five outcomes.[111]

How Do These Effects Compare with Findings from Previous Mentoring Studies?

Two meta-analyses of mentoring evaluations have provided excellent summaries of prior mentoring research. The first (DuBois et al. 2002) reviewed evaluations of programs published from 1970 through 1998, while the second (DuBois et al. 2011) examined programs evaluated from 1999 through 2010. These thorough reviews found that, on average, youth participating in mentoring programs have better emotional/psychological outcomes, feel more socially competent and have better academic/educational outcomes than youth who do not participate. These meta-analyses also showed that mentored youth exhibit fewer problem/high-risk behaviors—something we did not find in our study. Thus, the

Table 5.3
Comparison of Estimated Program Effect Sizes to Those of Recent Meta-Analyses

	Current Study: Random Assignment	Current Study: Quasi-Experimental	DuBois et al. (2011)	DuBois et al. (2002)
Emotional/psychological well-being	.14	.32	.15	.10
Social/relational outcomes[a]	.04	.10	.17	.15
Academic/educational outcomes[b]	.11	.05	.21	.11
Problem/high-risk behavior	.07	.04	-.21	-.20

Note: For purposes of this table, effect estimates for emotional/psychological well-being, social/relational outcomes and academic/educational outcomes are presented so that positive values reflect beneficial effects on outcomes; for problem/high-risk behavior, beneficial effects are reflected in negative values.

[a] Effect sizes for social/relational outcomes average those for parent trust, social acceptance and prosocial behavior.

[b] Effect sizes for academic/educational outcomes average those for grades, self-perceptions of academic abilities and skipping school. Skipping school is included in the academic/educational outcomes in this table to be consistent with how DuBois and colleagues categorized this measure in their meta-analyses. However, it falls under "academic behaviors" (a distinct group) in this study.

current study found improvements in some, but not all, of the areas that mentoring has previously been indicated to influence.

Compared with the effects found in DuBois et al.'s analyses, the current study found somewhat larger effects in the area of emotional/psychological well-being and smaller but still favorable effects in the areas of social/relational and academic/educational outcomes (see Table 5.3 on the previous page). However, unlike the meta-analyses we did not find effects of mentoring in reducing problem/high-risk behavior. In the next section, we explore how these outcomes might be affected by the different levels and types of risk that youth had experienced when they applied to the program.

Did Impacts Vary by Youth's Risk Profile?

As discussed in Chapter 2, the youth in this study faced a wide range of environmental and individual challenges and vulnerabilities. To assess whether program effects differed by the types and levels of risk youth faced, we examined the pattern of estimated program effects across the four risk groups described in previous chapters. To help ensure that our analyses provided the strongest chance to observe differences across these subgroups (that is, that our four risk profile groups were big enough to allow us to see patterns in the data), we used the full sample in these analyses, combining program participants from both the random assignment and quasi-experimental components of the study.

Table 5.4 on the next page shows the estimated program effects, or effect sizes, for each outcome and their statistical significance within each of the four risk subgroups, as well as, in the final row of the table, the average effect size across all outcomes for each risk group. Although we combined program participants from both mentored groups for these analyses, the sizes of the four subgroups are still substantially smaller than the size of the overall sample. Thus, the margins of error around the estimates are larger than those in Table 5.2, such that larger effects are required to reach statistical significance. The groups themselves also differed in size, so that those with fewer youth needed larger effects to achieve statistical significance. With these considerations in mind, we note not only statistically significant program effects *within* groups and significant differences

across groups, but also instances where effect size estimates suggested a *possible* effect (bolded in Table 5.4). We chose .15 for this threshold both because it is in the range of effects reported previously for youth mentoring programs and because this value was suggested as a cutoff for "small" effects based on an empirical analysis of intervention effect size estimates more generally (Lipsey 1990).

As shown in Table 5.4, we found a significant difference in estimated program effects across the four risk profile groups for three outcomes: parent trust, misconduct and the number of outcomes with meaningful positive change. For the remaining seven outcomes, significant differences were not found across risk groups. Thus, on the whole, we found only limited evidence that program effects differed across the four risk profiles.

Still, it is interesting to note patterns in the differences across groups and to examine those effects that reached statistical significance or at least our threshold for a "possible" effect (that is, .15 or greater) *within* each group. When doing so, we note a tendency for youth high on only individual risk to have gleaned the most benefits from program participation.[112] For two of the three outcomes whose effects vary across the risk groups—parent trust and the number of positive outcomes—the estimated program benefits were greatest for the group high only on individual risk. And this group, despite its relatively small size, also had the largest number of outcomes (three) for which effects reached the $p < .05$ level of statistical significance (a slightly more conservative criterion than the $p < .10$ level utilized throughout this report) as well as the largest number of outcomes (five) with favorable effects that reached or surpassed our .15 threshold. It was also comfortably ahead of the other groups in terms of the average size of all 10 estimated effects.

Several dynamics could help explain why this particular group of youth may have derived greater benefits from their program involvement. As discussed in Chapter 3, mentors want to feel needed. This motivation fits especially well with youth who have individual challenges, as these needs can serve as a clear focus for mentors' efforts. Youth, too, may be more receptive to mentoring when they have already experienced the negative consequences of engaging in problem behavior or struggling with issues in

Table 5.4
Estimated Program Effect Sizes for Different Risk Groups

Scale/Indicator	High Individual/ High Environmental ($n = 329$)	Low Individual/ High Environmental ($n = 295$)	High Individual/ Low Environmental ($n = 167$)	Low Individual/ Low Environmental ($n = 483$)	Significant Difference Between Risk Groups?
Depression	**-.19[†]**	**-.19[†]**	**-.36***	**-.19***	No
Parent trust	-.04	.03	**.38****	.10	Yes[a]
Social acceptance	**.19[†]**	.12	**.16**	.04	No
Self-perceptions of academic abilities	.14	.00	.12	.13	No
Grades	.09	.11	.14	.12	No
Skipping school	.12	**.20**	-.04	**-.27**	No
Misconduct	-.13	**-.17**	.13	**.36[†]**	Yes[a]
Prosocial behavior	.02	-.10	.00	.02	No
Number of outcomes showing positive change	**.16[†]**	**.35*****	**.25***	.03	Yes[b]
Number of outcomes showing negative change	.00	-.10	**-.22**	-.09	No
All outcomes	.08	.08	.15	.06	—[c]

Note: The analyses in this table combine the two mentored groups from the impact and quasi-experimental portions of the study. Standardized mean difference estimates of effect size are reported in the table. These effect estimates were derived from regressions that included the covariates listed in the text. Fixed effects were used to adjust for the fact that some of the youth appear in the dataset twice (once as a comparison group member and the next year as a participant group member). Effect sizes for outcomes that reached or exceeded the threshold of .15 (as discussed in the text), regardless of direction, are highlighted in bold.

*** $p < .001$ level of significance.

** $p < .01$ level of significance.

* $p < .05$ level of significance.

† $p < .10$ level of significance.

a For this outcome there was evidence of program effects differing as a function of environmental risk level (that is, a significant interaction of environmental risk status X participant- versus comparison-group status [$p < .10$]). The average program effect for both groups high on environmental risk differs from the average effect for both groups low on environmental risk.

b For this outcome there was evidence of program effects differing as a function of combined levels of individual and environmental risk (that is, a significant interaction of individual risk status X environmental risk status X participant- versus comparison-group status [$p < .10$]).

c We did not assess whether the average effect sizes across all outcomes differed as a function of youth risk.

their school performance or emotional well-being. At the same time, family support and other environmental resources (for example, in the youth's neighborhood) may be crucial to bolstering mentors' efforts and reinforcing youth gains. For youth high on environmental risk, such supports are more likely to be lacking. The youth's surrounding environment could be especially important when mentors are volunteers. Relative to helping professionals (like teachers or coaches), who often serve as "natural" mentors outside of formal mentoring programs, volunteers are likely to have much more limited training and experience working with youth (which was true for many of the mentors in our sample, as discussed in Appendix C)—thus, environmental supports may be particularly helpful in fostering these relationships.[113] At the same time, those mentors matched with youth high on only individual risk were the most experienced—that is, the most likely to have mentored previously (see Appendix C). We also discussed in Chapter 4 that these mentors reported the lengthiest support calls

and were more likely to have received early-match training and regular program support. The positive outcomes seen in this group may, in part, result from these differences in program practices and mentor experience. Finally, both groups of youth high on individual risk scored notably worse on our outcome measures at baseline,[114] thus giving them more room for improvement on these outcomes relative to youth in the other risk groups.[115]

Turning to the other three groups, although there were less clear differences among them, youth who were high on only environmental risk and those who were high on both types of risk appeared to show a more positive response to mentoring. For each of these subgroups, favorable program effects reached or approached significance for two important outcomes—depression and the number of outcomes with positive change—and reached significance or at least surpassed our threshold of .15 for one other outcome. In contrast, there was only one significant favorable program effect for the lowest-risk group (depression), despite this being the largest subgroup. For this group, there were also only two outcomes showing evidence of favorable program effects at the .15 threshold, and there was one outcome, misconduct, in which mentored youth showed significant *setbacks* relative to their peers.

Thus, our exploration provides hints that youth who had relatively low levels of both individual and environmental risk accrued the least benefits from mentoring. This pattern is only suggestive, however, and differs from findings in a recent meta-analysis that did not show a consistent pattern of weaker program effects when participating youth were relatively low on both individual and environmental risk, as opposed to high on one or both types of risk (DuBois et al. 2011).

Did Program Effects Vary by Other Youth Characteristics?

In addition to exploring the role of risk, we also wanted to explore whether other youth characteristics might be associated with bigger or smaller program benefits. We considered several subgroups that are interesting or "targetable" from a policy or program point of view (for instance, youth from single-parent homes) or that have been identified by prior research as being particularly receptive to mentors (such as youth who already have moderately good relationships with others).[116] If differential effects were found for any of these subgroups, programs might opt to target more of their resources toward groups for which mentoring is more effective. In particular, we examined the effects of the program for youth varying on the following characteristics at baseline (all based on parent reports except where noted):

- Gender;

- Race/ethnicity;[117]

- Age;

- Single-parent status;

- Living in poverty;

- Having a parent or other close family member who was incarcerated or had frequent problems with the law;

- Lacking a close friend;

- Experiencing potentially stressful life changes over the past year;

- Having poor, moderate or good relationships with parents and peers (based on youth report); and

- Having at least one "special" adult in their lives (based on youth report).[118]

In general, we did not detect differences in program effects as a function of these youth characteristics. (Results are summarized in Appendix B.) Thus, these findings as a whole do not suggest that programs could improve their effects by targeting youth based on any of these characteristics.

However, there were two sets of findings that are noteworthy. First, we found significant differences in estimated program effects for three outcomes—depression, parent trust and the number of outcomes showing negative change—based on whether youth were reported to have a parent or other close family member who was incarcerated or had frequent difficulties with the law. For each outcome, youth with such a family member (about one quarter of our sample) did *not* seem to benefit from mentoring, whereas significant program benefits were evident for youth who did not have a family member incarcerated. Having an incarcerated *parent*

in particular (which was likely the case for many of these youth) may be especially challenging as a context for mentors and programs to make significant inroads with youth (Eddy et al. in press). Clearly, though, even if program benefits are less consistent for this population of youth, a strong case can be made for seeking better approaches to serving such high-need youth rather than shifting away from working with them.[119]

The second set of findings involved differences in program effects for three outcomes—misconduct, grades and the number of outcomes showing positive change—based on the number of potentially stressful life changes experienced by youth during the previous year. These findings suggested more pronounced benefits of mentoring among youth who recently had experienced stressful changes or events. This pattern is consistent with the idea that mentoring is designed to provide youth with increased access to social support (Barrera, Bonds 2005) and thus may buffer the adverse effects of stress on youth well-being and adjustment (the "stress-buffering" hypothesis; see Cohen, Wills 1985). It should be noted, however, that the differences in program effects based on both having an incarcerated family member and stressful life changes/events were significant for only a minority of the outcomes examined. Thus, both sets of findings must be considered preliminary.

Summary

On average, after receiving almost 10 months of mentoring, participating youth fared better than similar youth without mentors, on a number of important dimensions. These included emotional/psychological well-being, social relationships, academic attitudes and self-reported grades. Reductions in symptoms of depression were particularly notable—both for the large size of the effect and the consistency across groups. Given the notable levels of depressive symptoms seen in this sample at baseline, the effect on depression is quite encouraging. Findings for our aggregate measure

of program outcomes further indicate that program participants were more likely to improve on at least one of our outcome measures and to show improvement on multiple outcomes than were youth in the comparison/control group.

This study has provided one of the first opportunities to look closely at how mentoring's effects vary by the types and levels of risk youth face. Overall, our results did not reveal a strong pattern of differences in effects across risk groups. However, there were some indications that youth high on individual but not environmental risk responded most favorably to mentoring. Thus, despite the distinct match challenges reported in Chapter 3 for these youth, they appear to have benefited slightly more than youth with other risk profiles. Youth in the other two "elevated-risk" profile groups (that is, high on both types of risk, or high only on environmental risk) also appeared to derive benefits that were somewhat greater than those in the lowest-risk group. Overall, these findings present an optimistic outlook for programs serving higher-risk youth. As we will discuss in the following chapter, if programs did more to provide tailored support to youth with differing profiles of risk, we might see even bigger and more consistent benefits for these youth.

The analyses reported in this chapter also revealed that—with two noteworthy exceptions (that is, slightly stronger effects for youth without an incarcerated family member and for those recently experiencing more stressful life changes)—program effects for the most part did not vary significantly depending on a range of other youth characteristics examined, such as their gender, race/ethnicity or age (within the age range of our sample—9 to 15 years old). Combined with our findings about youth risk, this aspect of our results is consistent with the conclusions of DuBois et al. (2011) that mentoring is a flexible intervention that can be effective for youth with a wide range of characteristics and backgrounds.

Conclusions

Chapter VI

T he goal of this evaluation was to assess how volunteer-centered, community-based one-to-one mentoring works for youth with varying levels and types and risk. What kind of mentoring relationships do youth experience? What challenges do their mentors face? Can mentoring benefit youth across the risk "continuum"? And, crucially, are there differences in youth's experiences or outcomes, based on their varying risk profiles? Finally, what might programs need to do to ensure that all types of youth develop strong, long-lasting and beneficial relationships with their mentor?

To address these questions, we conducted an evaluation involving more than 1,300 youth from seven different community-based mentoring programs. In addition to assessing youth outcomes, we collected extensive data about youth risk, the relationships that developed between the youth and their mentor, and the programmatic supports that youth, mentors and parents received. We created four groupings based on youth's risk levels in both environmental (that is, the surrounding environment) and individual (that is, the youth's behaviors or characteristics) areas: those relatively high on both types of risk; those relatively low on both types of risk; those relatively high only on environmental risk; and those relatively high only on individual risk. As we describe below, we found that mentors matched with youth in these four groups tended to report distinct training/support needs, challenges and reasons for match closure, which were largely in line with the different risks faced by the youth, supporting the usefulness of these particular groupings for future research and practice.

This is the first large-scale study to examine the role of youth risk in mentoring relationships and outcomes. More research will be needed to confirm our findings and conclusions. Yet, the patterns that emerged in this study are consistent across several areas (for example, challenges and match closures) and, in many respects, are in keeping with both prior research and practitioner observations.

Overall, we believe these results provide fairly compelling evidence that youth's risk profiles play a significant role in the development of relationships in community-based mentoring programs. The results also point toward several valuable lessons for practitioners and funders.

We addressed four sets of questions in this study:

(1) Can mentoring programs reach higher-risk youth? Are they already reaching such youth?

The seven programs in this study reached youth facing a wide range of challenges—without significant effort beyond their normal outreach strategies. Like many mentoring programs (most notably BBBS programs within the US), a large proportion of our sample of youth came from single-parent homes (about two thirds) and from poor households (over two fifths had annual incomes below \$20,000). In addition, more than two thirds faced some type of "individual-level" challenge, such as academic challenges or mental health difficulties, and almost a quarter had already engaged in problem behaviors, such as running away from home, bullying others, being suspended from school multiple times or being involved with a gang.

Compared with youth nationally, the youth in the study were more likely to have experienced many challenges—for example, they were almost twice as likely to live in a single-parent home, three times as likely to show serious signs of depression and several times as likely to be in foster care or to have experienced homelessness. About a quarter of the youth met evidence-based criteria proposed by Child Trends (Moore 2006; Moore et al. 2006) for "high socio-demographic risk." Few, however, were reported to have used drugs or to have already been involved with the juvenile justice system—factors that are well recognized as putting youth at very high risk for serious future problems. The low incidence of these experiences in our sample was perhaps due to the fact that the youth were fairly young (with an average age of a little over 11). For younger children, multiple school suspensions and bullying may be just as telling as drug use or arrests, since they, too, are associated with a greater likelihood of more serious problem behaviors down the road.

Taking all of these factors into consideration, the youth in our study are, on average, best categorized as "higher risk," rather than "high risk," though this simple designation ignores a range of considerations that can be important when characterizing youth's experiences of risk. Indeed, within our sample, we found substantial variability in both the levels and types of risk—individual and/or environmental—that youth faced. This variation was intentional and important, as it allowed us to explore how different risk profiles were associated with various match experiences and youth benefits.

(2) Can volunteer-centered, one-to-one community-based mentoring programs benefit higher-risk youth? Do benefits differ depending on youth's risk profile?

Findings suggest that, relative to the comparison group, mentored youth improved in their emotional/psychological well-being, peer relationships, academic attitudes and self-reported grades, but not in their behavior. About 13 months after their baseline assessment, when program participants had received an average of about 9.6 months of mentoring, they were doing better than youth in the non-mentored comparison group, on a number of important outcomes. In particular, these youth reported:

- Fewer depressive symptoms;

- Greater acceptance by their peers;

- More positive beliefs about their ability to succeed in school; and

- Better grades in school.

We also developed two new aggregate measures of the number of positive and negative changes that were evident for each youth across seven of our key outcomes: Program participants showed meaningful improvement in more outcome areas than youth who did not receive mentoring (but they did not show differences in the negative-change variable).[120] They were also more likely to show improvement in one outcome area and to show improvements in multiple areas. As noted, previous studies have found impacts for mentoring in multiple areas—yet these findings could reflect individual youth benefiting in multiple areas *or* distinct groups of youth

reaping discrete benefits (for instance, some youth making gains only in academics, with others benefiting only in peer relationships). The findings from the current study suggest that one of the strengths of mentoring may be its ability to benefit youth in both ways—that is, to benefit individual youth in at least one area that is hopefully tied to their specific needs and (for some youth) to promote gains across multiple areas in ways that foster overall trajectories of positive development.

However, we found *no* differences in youth's reports of:

- Their relationships with parents;

- Their positive ("prosocial") behavior toward peers;

- Skipping school; or

- Misconduct.

To make our final conclusions about program benefits, we considered findings from both the random assignment portion of the evaluation (in which youth from two programs were randomly assigned to receive mentoring, and their outcomes were compared with those of youth from the same two programs who were randomly assigned to a control group)[121] and the quasi-experimental portion of the evaluation (in which the outcomes of youth participants, drawn from all seven programs, were also compared with those of the control group—in this case called the "comparison group"). For *both* portions of the evaluation, findings indicated that mentoring reduced symptoms of depression and increased the overall number of outcome areas in which youth exhibited meaningful improvement. A parent-reported measure (available only for the random assignment sample) similarly indicated impacts on emotional/psychological well-being.

Improvements in youth's peer relationships, their views of their own academic abilities and their self-reported grades were found only for the quasi-experimental sample. Effects for these outcomes in the random assignment sample were positive, but notably smaller and did not reach statistical significance. The study thus provides more moderate evidence of benefits in these three areas.

Based on these findings, we offer the following conclusions about the benefits provided by the mentoring programs in this evaluation:

1. We found strong evidence that community-based mentoring improved youth's emotional/psychological well-being;

2. We also found strong evidence that mentoring increased the *number* of areas in which youth experienced meaningful improvements;

3. The study yielded more moderate evidence that youth improved in their peer relationships, academic attitudes and self-reported grades; and

4. We found no notable benefits in youth's misbehavior, skipping school or their relationships with parents.

The specific areas in which youth benefited are noteworthy, as they are important indicators of well-being and can facilitate broader improvements in youth's lives. Looked at another way, deficits in these areas can make children vulnerable to more serious problems down the road. For example, childhood depression is associated with suicidal behavior, academic and social difficulties, and increased risk for substance abuse and teen pregnancy (Cash, Bridge 2009; Malhotra, Das 2007)—a reality that is particularly sobering in light of the fairly high proportion of youth in this sample (nearly a quarter) who had serious symptoms of depression at baseline. Low peer acceptance and social isolation are similarly associated with dropping out, criminal behavior and health problems later in life (Parker, Asher 1987; Caspi et al. 2006).

The lack of detected benefits in other areas, particularly misconduct, differs from several earlier evaluations of youth mentoring programs and also differs from parent-reported impacts in this area within the study's random assignment sample. For example, findings from P/PV's impact study of BBBSA community-based mentoring as well as recent meta-analyses[122] suggest that mentoring can reduce conduct problems (DuBois et al. 2011; Tierney et al. 1995; Tolan et al. 2008) and promote gains in other behavioral outcomes such as school attendance (DuBois et al. 2011; Wheeler et al. 2010; Tierney et al. 1995).

The differences between our findings and those of past research could in part reflect the fact that the matches in this study were less than nine months long, on average. In the P/PV study, stronger program benefits were seen in longer matches, with effects on some outcomes (for example, drug use) not emerging until youth had received at least 12 months of mentoring (Grossman, Rhodes 2002). Also, in the present study, less than one fifth of youth spent time with their mentor three times a month or more, on a fairly regular basis (70 percent of months or greater), whereas in the P/PV study *nearly half* of the mentored youth reported meeting their mentor at least weekly. As more consistent and frequent mentor-youth contact has been linked to greater youth benefits (DuBois et al. 2002; Karcher 2005), the differences between the two studies in these areas could at least partially explain the divergent findings.[123]

As noted, parent reports did suggest significant improvements in youth behavior. However, we did not weigh these reports in our overall conclusions about youth benefits because we only collected this measure for the random assignment sample—thus we could not be sure that we would have seen similarly positive effects in the quasi-experimental component of the study. Yet, in this case, parent reports may have actually provided a more sensitive measure of misconduct in this fairly young sample. Parents rated youth on their frequency of engaging in very basic types of misbehavior that are not uncommon for both younger and older youth—for example, fighting with or bullying other children, losing their temper often, lying, cheating or stealing things. Although several items do overlap with our youth-reported misconduct measure, that measure addressed a wider range of behaviors, some of which are much more common among older youth—for example, carrying a weapon or being a member of a gang.

Overall, benefits were fairly similar for youth regardless of their risk profile, with some notable exceptions. As reported in Chapter 5, our findings suggest program benefits did not vary markedly as a function of the risk backgrounds of youth. Indeed, effect sizes indicative of "possible" benefits (.15 or higher) were apparent for at least two outcomes (typically more) for each of the four risk groups, suggesting that youth derived some gains from

participation regardless of risk status. Moreover, for one specific outcome—depression—program effects reached or approached statistical significance in all four groups. Similarly, effects on our aggregate measure of positive change reached or approached statistical significance in all but one of the four groups. This pattern suggests that the benefits of volunteer-centered community-based mentoring are not confined to youth with particular types or levels of risk exposure.

As was also described in Chapter 5, there were some exceptions to this general pattern. Most noteworthy is the trend toward larger program benefits for youth who were high on individual, but not environmental, risk.[124] When considering effects that were either statistically significant or approaching significance ("likely benefits," as outlined in Table 6.1 on the next page) or not significant, but non-trivial in magnitude (.15 or larger, "possible benefits"), this group seemed to experience the biggest gains across several measures. This group was also the only one to experience significant gains in parent trust. Perhaps individual challenges like those experienced in this group (for instance, academic problems, depressive symptoms) provide a focus for mentors' efforts and help them feel needed. At the same time, family support may be crucial to bolster the mentoring relationship and reinforce youth gains. Families under stress in high-risk environments may be less able to provide this kind of support.

Of further note is our finding that mentoring most clearly fostered a greater number of meaningful gains on outcomes (as indexed by our "aggregate" measure of positive change) when youth were high on only one type of risk (individual or environmental).[125] Youth high on both types of risk received significant benefits, albeit not as large as these two groups, whereas the lowest-risk group did not experience measurable benefits on this outcome. Youth without significant levels of risk may not need help in as many areas and thus may experience fewer gains. Therefore, programs may be most successful when youth present mentors with moderate (but not absent or overwhelming) levels of challenge (DuBois et al. 2011). In line with this possibility, Schwartz et al. (2011) found that youth who entered BBBSA school-based mentoring programs exhibiting moderate levels of relational difficulties

benefited more from the program than youth for whom such difficulties were either severe or relatively absent.

(3) What is the quality of these matches? Do the characteristics of the match differ based on youth's risk profile?

Mentors and youth reported fairly strong relationships, regardless of youth's risk profile, and 53 percent of matches lasted beyond a year. We explored three aspects of youth-reported relationship quality: closeness (reflecting youth's general feelings of connection with the mentor); growth (the extent to which the relationship focused on achieving goals and youth growth); and youth centeredness (the extent to which the relationship focused on the youth's interests—for example, the activities youth wanted to engage in). Youth reports of the three dimensions of relationship quality, and mentors' reports of closeness, on average, were all fairly high, with almost three quarters of youth reporting at least a moderately positive relationship with their mentor across all three dimensions, and these reports did not vary to a great extent across youth with different risk profiles. Match length, meeting frequency and total "dosage" of mentoring (that is, the number of hours the match met in total throughout the study period) were also fairly consistent across the risk groupings.

Mentors were responsible for closing more than half of the matches, as opposed to the parent, the program or the youth, each of which contributed to less than a quarter of closures. This finding (based on mentor reports) highlights the importance of ensuring that mentors are prepared, supported and satisfied with the match. The most common reason for match closures (cited by a third of mentors) was not enough youth interest. Not all youth are ready for, or want, a mentor. Programs need to assess why youth want a mentor (not just their parents) and screen out youth who are not likely to be active participants in the relationship. Along similar lines, 18 percent of mentors reported that the match closed because the youth did not seem to need a mentor. In contrast, only 10 percent reported that the match closed because youth's needs were too severe. Volunteers want to feel needed and that their time makes a difference for the child they're matched with, despite the challenges that youth with more serious needs may bring to the match.

Table 6.1
Summary of Results for the Four Risk Profile Groups

Low Environmental Risk	
Low Individual Risk	**High Individual Risk**

Youth Characteristics	Youth Characteristics
32% lived in poverty 58% lived in a single-parent home 6% had experienced homelessness 15% had a close family member who was incarcerated or often in trouble with police 4% had been suspended from school 2+ times in the last year 18% showed serious signs of depression 1% had police contact	23% lived in poverty 52% lived in a single-parent home 6% had experienced homelessness 18% had a close family member who was incarcerated or often in trouble with police 19% had been suspended from school 2+ times in the last year 24% showed serious signs of depression 9% had police contact
Benefits	**Benefits**

Likely[a]
   Depression
   (Misconduct worsened)
Possible[b]
   Skipping school
	Likely
	   Depression
	   Parent trust
	   Number of outcomes with positive change
	Possible
	   Social acceptance
	   Number of outcomes with negative change

Match Characteristics	**Match Characteristics**

Expectations
More likely to have unmet expectations in…
   Time commitment
Meetings/Match Length
NA[c]
Training/Support
NA
Challenges
More likely to report challenges in…
   Conversations
Closure
More likely to close…
   Due to low youth interest
   Because youth didn't seem to need a mentor
	Expectations
	NA
	Meetings/Match Length
	   Less rematching
	Training/Support
	Relatively high training needs in…
	   Youth's emotional needs
	   Youth's social needs
	Reported longer support calls
	More likely to receive regular support
	More likely to receive early-match training
	More likely to experience enhancements
	Challenges
	More likely to report challenges in…
	   Managing behavior problems
	Closures
	More likely to close…
	   At program's request
	   Because youth didn't seem to need a mentor

Note: This table summarizes the characteristics of youth in each risk group. It also highlights the areas in which findings suggest "likely" or "possible" benefits of mentoring for the youth in each risk profile group. Finally, the table notes the distinctive features of their mentoring relationships (for example, match length, challenges), relative to those experienced by youth in the other groups. Thus, for example, the closure reasons listed for each profile are those in which the mentors were more likely (relative to those paired with youth in one or more of the other risk profile groups) to report that particular reason for closure.

[a] "Likely" benefits are those that were significant or marginally significant for the risk profile group involved (that is, $p < .10$ or better).

[b] "Possible" benefits are those that did not reach or approach significance (that is, $p \geq .10$), but for which estimated effect sizes were in a favorable direction and non-trivial in magnitude (that is, .15 or greater).

[c] There were no significant differences from other groups on this characteristic.

The Role of Risk: Mentoring Experiences and Outcomes for Youth with Varying Risk Profiles · *Conclusions* 67

Table 6.1
Summary of Results for the Four Risk Profile Groups, *continued*

High Environmental Risk	
Low Individual Risk	**High Individual Risk**
Youth Characteristics	**Youth Characteristics**
60% lived in poverty 80% lived in a single-parent home 21% had experienced homelessness 29% had a close family member who was incarcerated or often in trouble with police 4% had been suspended from school 2+ times in the last year 18% showed serious signs of depression 1% had police contact	54% lived in poverty 71% lived in a single-parent home 26% had experienced homelessness 35% had a close family member who was incarcerated or often in trouble with police 28% had been suspended from school 2+ times in the last year 32% showed serious signs of depression 15% had police contact
Benefits	**Benefits**

Likely
    Depression
    Number of outcomes with positive change
Possible
    Misconduct
    (Skipping school worsened)
	Likely
	    Depression
	    Social acceptance
	    Number of outcomes with positive change
	Possible
	    None

Match Characteristics	**Match Characteristics**

Mentor Expectations
More likely to have unmet expectations in…
    Family's needs
Meetings/Match Length
    More youth cancellations
Training/Support
More likely to receive ongoing training
Challenges
More likely to report challenges in…
    Youth's preparation for meetings
    Family support
    Family asks for too much help
Closure Reasons
NA
	Mentor Expectations
	More likely to have unmet expectations in…
	    Youth needs
	Meetings/Match Length
	    More youth cancellations
	    More character/behavior change activities
	    More growth/goal focus
	Training/Support
	More consistent support
	More hours of ongoing training
	Relatively high training needs in…
	    Interacting with youth's family
	    Social services
	    Youth's emotional needs
	    Youth's social needs
	    Youth's behavioral needs
	    High-risk youth
	Challenges
	More likely to report challenges in…
	    Conversations
	    Youth's preparation for meetings
	    Family support
	    Family asks for too much help
	    Managing behavior problems
	    Bridging economic differences
	Closure Reasons
	More likely to close…
	    Because youth needs were too severe
	    Due to differences in interests or personalities

At the same time, it is important to note that these findings reflect *mentors'* interpretations of why the match closed, and their perceptions of a lack of need (or interest) on the part of youth may not always be accurate. This underscores the importance of helping mentors understand that youth may come to a relationship with a wide range of needs, some of which may not fit mentor's preconceptions (a theme we discuss in more depth below).

The similarities in relationship quality and duration across the risk subgroups belie very different challenges and reasons why matches ultimately closed. (See Table 6.1.) Our measures of relationship quality and length (two of the most frequently used gauges of relationship "success" in the field) may go only so far in describing the day-to-day interactions in these matches. The struggles faced by mentors and why their relationships ultimately closed depended in large part on the risk profile of the youth with whom they were matched. Specifically, our findings suggest the following:

- **Youth <u>low on both</u> environmental and individual risk (our "lowest-risk" group):** Only about a third of these youth lived in poverty, and 58 percent lived in single-parent homes. Matches involving these youth were relatively more likely to close because the youth did not appear to be interested in the relationship. Mentors were also more likely to report difficulties in having conversations with these youth (perhaps due to their perceived lack of interest).

- **Youth high on environmental and low on individual risk:** The vast majority of these youth lived in single-parent homes, and three fifths lived in poverty. Additionally, 29 percent had a close family member who was incarcerated or often in trouble with the law. Mentors of these youth reported more youth cancellations of meetings and more difficulties related to a lack of family support for the match as well as intensive family needs.

- **Youth high on individual and low on environmental risk:** Only about a quarter of these youth lived in poverty, and a little more than half lived in single-parent homes. However, almost a quarter showed serious signs of depression, and 19 percent had frequent school suspensions. These

matches were relatively more likely to close at the program's request, and this group of youth was the least likely to be rematched. Mentors of these youth faced challenging youth behaviors and reported needing training and support around this issue. But, importantly, as noted above, this group received relatively strong program benefits.

- **Youth <u>high on both</u> environmental and individual risk (our "highest-risk" group):** This group represents what is commonly thought of as "high-risk" youth. Almost three quarters lived in single-parent homes, and more than half lived in poverty. About a quarter had experienced homelessness in the last five years. Almost a third showed serious signs of depression, more than a quarter had frequent school suspensions, and 15 percent had had recent contact with the police or had been sent to juvenile hall. Notably, more than a third of these youth had a close family member who was incarcerated or often in trouble with the law. These youth reported a relatively greater goal or "growth" focus to their relationship. Mentors similarly reported engaging in more character- or behavior-change activities with these youth. Mentors in these matches were relatively more likely to report a whole host of challenges; not surprisingly, these related to both the youth's family circumstances, as well as interactions with the youth him or herself. They also reported multiple training and support needs reflecting these many challenges.

The risk profile of the youth in this group seems likely to be the higher end of what the typical volunteer mentor can handle (at least at the level of training and support provided by these programs). Such youth are struggling *both* with challenges in their surrounding environments and with their own behaviors or mental health issues—either of which can be quite difficult for a mentor to work with. Together they are likely to be all the more formidable. Yet, importantly, the mentors of these youth were the *most* likely to say they would consider mentoring again (and these youth did benefit from the program). This finding affirms the value mentors place on feeling that youth truly need them—despite the reality that such young people may bring far greater challenges to the relationship.

(4) What kinds of practices and programmatic supports are needed to enable mentoring to succeed with youth of varying risk profiles?

Youth Enrollment:

In addition to more careful screening of youth to ensure that they do in fact want a mentor, our study also suggests the value of measuring youth risk at intake—both the environmental and individual challenges youth may be experiencing—for example, through a brief parent questionnaire. Youth's experiences in both of these areas may have important implications for the match's success—the specific types of challenges mentors are most likely to experience and the types of supports that may be most helpful.

Training and Support:

Both training and support were linked with match strength and length. Programs reported that only about half of the mentors in our study received pre- or "early-match" training. Mentors who received this training met more frequently with their mentee, had matches that were more likely to last at least 12 months and were matched with youth who reported higher levels of relationship quality. Ongoing training was also linked with longer, higher-quality matches.

Early-match (and ongoing) training—particularly the strict definition of "early-match training" used in this initiative—likely provides valuable tools for guiding mentors in how to approach their mentoring relationship; and all mentors should have at least some orientation or training prior to starting their match. Whether mentors in this study took advantage of training (and other available supports) also probably says something about the mentor him or herself. Mentor's receipt of early-match training, for example, likely depended on mentor characteristics (such as time available, initiative and seriousness about the program) that may have shaped how they approached the relationship with or without the training—and it may be these characteristics that determined both the match's success and whether the mentor got the training. Nevertheless, programs can still use mentors' receipt of early-match training to help predict match outcomes. Mentors who don't attend training may need additional supports to provide the information they would have gotten in training and/or to offset whatever might have led them to not participate.

Regular support calls to mentors were also linked with match success, although not as strongly as early-match training. Mentors who received match support (typically in the form of case manager's support calls) during at least 70 percent of their months in the program met more frequently and had longer-lasting matches than mentors in matches with less consistent support. However, they did not have higher-quality matches (according to youth reports). Additionally, matches that received the program enhancements (a combination of early-match training and regular support contacts for mentors)[126] were higher quality, longer lasting and met more frequently than those that did not. Support contacts with parents and youth did not appear to be as central to match success; they were linked only with the frequency of match meetings.

These findings provide good preliminary evidence that program support may help strengthen mentoring relationships and promote match success. However, the very challenges that make it difficult to reach families and mentors with support calls (for example, a lack of interest, higher stress levels) may also interfere with the match itself. Thus, although our analyses suggest that support may contribute to stronger and longer matches, they cannot answer the question of whether support *causes* match success. This is clearly an important area for future research.

Matches with youth of different risk profiles experienced similar types and levels of support, with one important exception. Generally, we did not find strong differences across risk profiles in the mentor's receipt of training and support. However, mentors matched with youth who were high on individual risk appeared to receive more intensive support than others. For example, mentors matched with youth who were high on both individual and environmental risk received more hours of ongoing training and more consistent support from program staff. And those matched with youth high on only individual risk were more likely to have received early-match training as well as more

consistent and longer support calls. These mentors were also relatively more likely to have started their match with previous mentoring experience. It is intriguing that this latter group of youth (high on only individual risk) also appeared to reap relatively strong benefits from their program participation. The differences in practice for this group of youth suggest that it may have been a combination of youth characteristics, mentor characteristics and strong programmatic support that helped this group benefit.

We also wanted to know whether training and support might be particularly important for youth of varying risk profiles, addressing the question: Are program practices related to match outcomes to different degrees depending on youth risk? We did not find any strong patterns in the data to support this idea.

Managing the youth's family was often challenging, particularly for mentors matched with youth high on environmental risk. This finding points to the importance of training mentors about what to expect and of orienting *families* up front to their roles and responsibilities. Many of these families face significant hardship. Yet, it is important that they understand the mentor's value—and that they have a role to play in making the mentor feel welcome and supported in their home. Programs should try to ensure that families buy into the match from the start and should communicate to all parties that it takes time for a child and an adult to "click." Parents can be crucial sources of support by urging youth (and mentors) to continue with the relationship and by sharing information with mentors about their child's perceptions and progress. In response to these issues, some programs have started to require that parents attend an orientation or training session prior to the enrollment process being initiated. In this study, only one of the seven programs had pre-match training for parents when the study began; four offered orientation; and two offered neither. By the end of the study, one of the programs had added parent orientation to its practices.

Many mentors were surprised by the needs they encountered in their match. About a quarter of mentors reported that their expectations about youth's needs had not been accurate, while almost

two fifths said their expectations regarding the youth's family's needs had been wrong or incomplete. Several mentors similarly noted in their responses to our open-ended survey question that they started their relationship with very different expectations than the reality with which they were confronted once matched. This highlights the necessity of preparing mentors, as much as possible, for the individual youth with whom they will be paired. In most cases, mentors felt that staff were forthcoming with them. Yet, sometimes the staff themselves lacked extensive (or accurate) information about the youth's needs—this again points to the importance of collecting in-depth information about the youth *and* his or her family during enrollment and of training mentors to "expect the unexpected." By impressing on mentors that unanticipated challenges *will* arise and preparing them to deal with such challenges (for example, how to help families access appropriate resources), programs may increase the odds that youth are connected with the right kinds of support and matches are sustained over time.

In some cases, however, mentors' unmet expectations did not result from youth having more severe needs than the mentors felt prepared to deal with, but rather, they appeared to stem from mentors feeling that the youth simply did not need them. Volunteers may not always understand the wide variety of needs that can predispose a child to poor outcomes. For example, while most of the youth in the study were clearly experiencing economic challenges, less than a third of mentors reported that "poverty" was an area of need for their mentee. This belief may simply suggest that mentors did not feel that poverty stood out as a particularly large problem for the child they were matched with. However, it also may have resulted, in part, from the fact that many of the volunteers had very little personal experience with the types of issues these youth faced. Two fifths of mentors reported that they had not faced (in their own histories) *any* of the challenges we asked about in our survey (which included poverty, family struggles, school challenges and problems with parent or peer relationships). Only 12 percent specifically reported having experienced poverty. And although 40 percent reported past professional experience with youth (for example, as a teacher or youth worker), and a substantial number had experience with youth

facing a range of risk factors, almost a third had *no* experience with youth facing any of the risk factors we outlined. This suggests that many mentors may have started their mentoring relationships with very little foundation for understanding the issues that youth were dealing with. Mentors may need to be educated, for instance, about the varied ways that poverty can affect youth's lives (such as a lack of consistent phone service, different constraints on time, heightened stress levels, etc.) and how it can result in a family that operates very differently from their own. Poverty can be just as detrimental a force in youth's lives as individual risk factors. This knowledge could help mentors understand how much they are needed across a wide variety of youth backgrounds.

Staffing:

The level and quality of case manager support may help determine a match's quality and longevity. Corroborating "program wisdom," which holds that the quality of case manager support can affect the strength and length of a mentoring relationship, we found that supervisors' ratings of case managers' competence were associated with mentor reports of the quality of the support they got from the program. Mentors' experiences of support, in turn, predicted match length as well as youth's reports of the quality of the relationship. These findings provide some of the first quantitative evidence that case managers can, in fact, influence the strength and length of matches in mentoring programs. Programs could build on this finding by taking steps to ensure that their hiring, training and supervision practices put case managers in the best position to effectively support matches.

Rematching:

The findings in this study do not warrant strong conclusions or a "bottom line" about whether rematching is an effective program strategy for youth when their first match ends. Yet, we found hints that rematching should be approached with caution. Namely, among youth who had been rematched, program effects for certain outcomes appeared to dissipate with greater length of time mentored. It is possible that some youth are just not good candidates for the mentoring that can be made available through a particular program, at least at a particular stage in their development.

Further, should youth—even due simply to the "luck of the draw"—be faced with a second disappointing experience with a mentor, there may be heightened potential for internalizing that loss in ways that harm youth (for example, lowered self-esteem). At a minimum, it would seem prudent for programs to review processes in place, to ensure they give these possibilities due weight in decisions about rematching.

Implications for Practitioners and Funders

The findings from this study have several notable implications for practitioners and funders:

- **Training and support for matches should be tailored to the types and levels of risk experienced by youth.** In general, mentors in this study felt fairly well supported and trained. However, different mentors experienced different challenges—depending at least in part on the risk characteristics their mentee faced. And although matches involving higher-risk youth seemed to present greater challenges, it is clear that even those with the lowest levels of observed risk brought their own distinct issues and concerns. This highlights the need to tailor program training and support to the specific levels and types of risk faced by participating youth—both the general population of youth served by any given program and the individual youth within each match.

  To do so, programs will need to systematically assess the risk profiles of participating youth (as noted above). These data could then be used to ensure that a program's practices are responsive to the full range of risk backgrounds of participating youth. In addition, programs can tailor training and support around the specific types and levels of risk experienced by *individual* youth—an approach that has shown promising results in related fields, such as education and health promotion. Mentors of youth who are high on environmental risk, for example, may need tailored support around understanding how to work with youth who may be more likely to miss meetings, have disconnected phones or have parents who may not be as available to support the match. And mentors who are matched with youth who are low on both types of risk may benefit from training on how to engage youth who may not *appear* to be

interested. Similarly, programs should take time with all new volunteers to discuss the background of their mentee and his or her family to help mentors understand why they are being matched with the youth and the specific needs they could help to meet—even when these needs may not be immediately apparent. Funders should support programs' efforts to better measure youth risk and to tailor the training and support they offer accordingly.

- **Mentoring should be broadly available, as youth with varying levels and types of risk appear to derive important benefits.** Overall, despite fairly extensive analyses, the study did not find strong evidence that mentoring benefited youth differently based on their risk profile or other background characteristics. These findings argue against restricting eligibility (or unduly weighting recruitment efforts) to youth with particular risk profiles or backgrounds, at least for programs that are structured similarly to the ones in this study. At the same time, for programs interested in targeting higher-risk youth (for example, in response to stakeholder priorities or evidence that benefits for a particular group would be especially valuable from a societal or policy standpoint), the study's findings provide optimism that such youth can be recruited and that, with the right supports in place, they can derive significant benefits from mentoring.

- **Greater emphasis should be placed on the mental health needs of youth and the benefits that mentoring can provide in this area.** Depression has been linked to a host of short- and long-term problems for young people, including suicidal behavior, academic and social difficulties, and increased risk for substance abuse and teen pregnancy (see Cash, Bridge 2009; Malhotra, Das 2007). It was striking that almost one in four youth in this study reported high levels of depressive symptoms at baseline. Our findings offer robust evidence that participation in volunteer-centered one-to-one mentoring programs can ameliorate and/or prevent the emergence of depressive symptoms. This is highly encouraging, given the number of other areas (personal, social and academic) that may benefit from better mental health. One key implication for programs is the importance of careful screening for mental health issues, both at intake and over the course

of a young person's involvement in the program, in combination with appropriate referral mechanisms for youth who are in need of additional support. At the funding level, the findings from this study suggest that mental health outcomes should be given greater weight in designing and evaluating the success of mentoring initiatives.

- **Efforts should continue to improve the strength and consistency of the benefits youth derive from mentoring programs.** As a whole, the findings of this study point to a positive, but not entirely consistent pattern of benefits for youth who had access to volunteer-centered, one-to-one community-based mentoring over a 13-month period. For example, the random assignment portion of the study provided limited evidence of significant program effects, with clear gains on only two of the measures that were the focus of our analyses, and neither portion of the evaluation showed that mentoring helped to curb involvement in problem behavior. These aspects of the study's results underscore a need for moderation when forecasting the likely impact of mentoring interventions (see Wheeler et al. 2010; DuBois et al. 2011). The findings also suggest, however, that by improving program supports (such as the training provided to mentors or to the staff who support the matches), it may be possible to strengthen the quality and longevity of mentoring relationships and potentially, in turn, increase the impact of mentoring programs on youth outcomes. Funding support will be necessary to make large-scale inroads in this area. These efforts should include support for intermediary organizations that can broker needed technical assistance and bring programs together to share lessons about effective practice.

In addition to these implications for practitioners and funders, the study points to several important areas for future research. For instance, studies should continue to examine links between youth risk and program effectiveness—particularly, whether tailoring program practices to the risk profiles of individual youth can improve mentor and youth experiences and, ultimately, increase program benefits. In doing so, researchers should develop more rigorous and reliable methods to capture program effects that appear in different outcomes for different youth—using, as a starting

point, the aggregate measures developed for this study. Without measures that are sensitive to more individualized patterns of change in outcomes, the benefits of mentoring program participation may be seriously underestimated. Clearly, too, we need to learn more about the role of mentoring in helping youth avoid depression and, in turn, potentially even more serious outcomes (for example, suicidal behavior). Finally, an important next step for both research and practice will be to consider youth's strengths and assets as part of their overall profile. Such resources—whether in the youth's environment or part of their personal make-up—can undoubtedly play a key (and perhaps complementary) role to that of risk in shaping match experiences and the benefits that accrue to youth.

Overall, we believe the findings from this study present an optimistic picture of the role that mentoring programs can play for youth facing a wide variety of risks. It is notable in this regard that youth in all four risk subgroups showed evidence of program benefits, including mental health gains that have the potential to lead to improvements in other areas. This was true even for youth who are often deemed "hardest to serve" in social programs—that is, those who were relatively high on *both* environmental and individual risk at intake. If programs offered more targeted training and support for matches involving higher-risk youth, they might be able to build on this foundation to foster greater benefits—helping more of these young people realize their potential and putting them on a path toward a stable, productive future. In sum, the high hopes that policymakers and funders have had for mentoring programs serving higher-risk youth may indeed be well founded, particularly if programs continue to refine their efforts to ensure that matches get the types of support that this study suggests can make a positive difference.

Endnotes

1. For meta-analytic reviews of research examining the effectiveness of mentoring programs, see DuBois et al. 2002, 2011.

2. The contents of this report were subject to the approval of the Bill & Melinda Gates Foundation.

3. One program was based in Oregon, but agreed that at least 80 percent of participating youth would live in adjacent Washington communities that also were part of its service area.

4. To select programs for involvement in the project, WSM issued a request for proposals to 11 programs that appeared to meet the criteria outlined in the Requirements for Programs in the Initiative text box. Initial phone calls disqualified three of these programs because they did not meet all of these criteria. The remaining eight programs, plus one additional program, submitted proposals. One of these nine programs was disqualified after submitting its proposal (again because it did not meet all of these criteria). WSM conducted site visits to the remaining eight programs to explore practices, capacity and sustainability. Each agency was given a score on its capacity and implementation of the Elements of Effective Practice. Seven were selected for involvement in the evaluation based on their ability to deliver high-quality services as evidenced by past performance; capacity to add a significant number of new matches to their programs in a relatively short time period; ability to sustain match slots beyond the three years of foundation funding; commitment to devote staff time and effort to work closely with P/PV to collect data for the study; experience working with outside evaluators; and clarity in key components of the proposal, including timeline, performance benchmarks, key personnel, project description and budget.

5. During the second and third years of the study (after completing their follow-up survey), the youth in the comparison group from the two largest programs could be matched if they still wanted a mentor. A total of 231 were ultimately matched. The changes they experienced during the second 13-month period are included in the quasi-experimental participant group's data (consequently, the 160 of these 231 youth for whom we were able to collect follow-up data are among the 615 youth in the participant group from the quasi-experimental portion of the study). Fixed individual effects modeling was used to account for the fact that youth could be in the dataset more than once. More details on this are available in Appendix B.

6. In other words, the comparison group youth all resided in two of the seven program sites and all applied to their programs mainly in 2008, while youth in the participant group were enrolled in both 2008 and 2009.

7. One of the main weaknesses of most comparison-group studies is that youth who enroll in a program often differ from those who did not apply. For example, applicants may be more motivated than nonapplicants. Thus, even without the program, the participant group's outcomes might have improved more over time than those of nonapplicants. This self-selection issue is much less of a problem with our study design because all of our comparison-group youth applied to receive a mentor.

8. We statistically adjusted for observed differences between the two groups; however, because we could adjust only for differences we measured, we cannot be completely sure, even with these adjustments, that the comparison group provides an accurate picture of how the participant youth would have done without mentoring.

9. The number of youth likely to benefit from mentoring was calculated based on estimates by Dryfoos (1990) that roughly one half of young people in the US are "at risk" for future negative behaviors, based on their demographic characteristics, their attitudes and behavior, and the social and economic context that surrounds them. Census data in 2000 indicated that 35.2 million youth were between the ages of 10 and 18 (half is 17.6 million). More specifically, MENTOR recommended targeting youth for mentoring services who are at "moderate" or "high" risk using Dryfoos' definitions (about 40 percent of all youth) as well as some youth at "low risk" (about half of all youth). Those youth at "very high risk" (about 10 percent of young people) were likely in need of multiple interventions. Mentoring was thus inferred to be less effective with this group so it was not included in the total estimate (see www.mentoring.org/about_mentor/value_of_mentoring/closing_the_mentoring_gap for more information).

10. Establishing reliable weights for risk factors appropriate to their relative importance, or severity, is challenging for a number of reasons, including the likelihood that such weights will vary according to the type of outcome being predicted, as well as age, gender and other characteristics of the youth involved. For these reasons, we did not attempt to incorporate differences in risk factor severity into our risk classifications.

11. Some research (for example, Sameroff et al. 1998) has suggested "tipping points" past which an increase in the number of risk factors is notably more predictive of difficulties for youth than are differences at lower levels of risk factor exposure. However, findings in this area have not been consistent (for a review, see Gerard, Buehler 2004).

12. It should be kept in mind that the distinction between individual versus environmental risk factors is itself somewhat of a simplification given that many sources of risk, such as difficulties with peers, likely reflect contributions of both the youth and his or her environment.

13. In addition to the analyses presented in this chapter, we also assessed whether various demographic groups were more or less likely to be higher risk. Specifically, these analyses looked at whether youth's gender, age and race/ethnicity were associated with youth's risk profile controlling for the other demographic variables as well as the mentoring program to which the youth applied. As expected, older youth were more likely to experience individual risk factors than younger youth (as such, youth in both "high individual" risk categories were significantly older than youth in both "low individual" risk

categories). The average (model-estimated) age of youth in the highest-risk group was 11.57; 11.69 for youth high on only individual risk; 11.19 for youth high on only environmental risk; and 11.02 for youth in the lowest-risk group. Even after accounting for age differences, African American youth were disproportionately represented in the low (versus high) individual risk categories. Proportions of African American youth in each group were as follows: 23 percent for both the lowest-risk group and the group that was high on only environmental risk; 16 percent for the group that was high on only individual risk; and 19 percent for the highest-risk group. Proportions in the first two of these categories were significantly higher than the third. Finally, Hispanic youth were significantly more likely to be in the group that was high only on individual risk (25 percent of the youth in this category were Hispanic) than in the highest-risk category (17 percent of these youth were Hispanic). The last two categories (both 21 percent Hispanic) did not differ significantly from the others in the proportion of youth who were Hispanic.

14. The questions we asked parents were slightly different from those used to create the national statistics to which we had access. We note differences for each measure where relevant.

15. We did not have access to a direct measure of whether the family was able to buy or own a home. Thus, to indicate the presence of this risk factor, we used parent reports of living in public housing or living in a situation that was not stable (that is, the parent believed it might be necessary to leave). To estimate the number of siblings living in the home, we took the total number of people reported to live in the home minus the number of different types of adults (for example, mother, grandmother) that were reported to be living with the youth. In addition, we did not ask about the education levels of both parents (in cases where the child lived with *both* parents), only the parent who completed our survey. Thus, we could include education level only of this parent (in most cases, the child's mother) in this measure.

16. These data are available, of course, only for youth in one of the two groups that had access to mentors. For youth who had more than one mentor, we report responses from the last mentor the youth had who completed a mentor follow-up survey.

17. In the P/PV study, about half (49 percent) of youth who were 9 to 14 years of age sustained their relationship 11 months or longer, whereas only about a third (34 percent) of youth 15 or over sustained their relationship that long. Similarly, whereas close to half (47 percent) of youth with no arrests sustained their relationship 11 or more months, 39 percent of those with one arrest and less than a quarter (23 percent) of those with two or more arrests had relationships that long.

18. The findings from the mentor survey reported in this chapter are based on the survey responses of 915 mentors who were the youth's most recent mentor at follow-up. These are the mentors whom youth reported on in their surveys. Limiting analyses to these mentors thus ensured that mentor and youth data pertained to the same relationships.

19. In all of these analyses, we controlled for (or held constant) the following variables: youth race/ethnicity, gender and age; program in which the youth was served; and, if the measure was derived from the youth survey at follow-up, mode of administration of this survey (that is, phone or in-person). Only findings (comparisons) that reached a significance level of at least $p < .10$ are reported here and throughout the report.

20. One of the enhancements implemented at the two agencies (only during the randomized trial portion of the evaluation) was to ask matches to meet in person at least three times a month; reciprocated telephone or email communication could substitute for the fourth in-person meeting.

21. The percentage of matches that met at least three times monthly 70 percent or more of the time was, as would be expected, significantly greater when the program expectation was for matches to meet at least this often (18 percent) as opposed to only at least two times a month (11 percent).

22. Note, however, that when we combined the frequency and duration of match meetings we did see one difference across risk groups. Almost two fifths of the matches (37 percent) met, on average, at least twice per month for a total of at least six hours. This level of meeting was true significantly more often for matches with youth who were low on both types of risk (40 percent) than it was for those with youth who were high on both types of risk (30 percent) or high on only environmental risk (33 percent).

23. About a fifth (21 percent) of mentors matched with youth who were high on both types of risk and 22 percent of those matched with youth who were high on only environmental risk reported frequent meeting cancellations. Mentors in these groups were significantly more likely to report frequent meeting cancellations than mentors matched with the lowest-risk youth (13 percent).

24. We also examined which one of these groups of activities was reported as the most frequently engaged in for each mentor. For the vast majority—86 percent of mentors—"fun" was the most frequent, followed by talking (10 percent). Character or behavioral change was reported as the most frequent activity only for 4 percent of mentors, and enrichment and structured activities were both reported as most frequent by less than 1 percent of mentors.

25. A total of 33 of the 36 mentors who chose this option were from the university-based program.

26. Mentors could indicate more than one factor as leading them to choose their specific focus.

27. Almost three fifths (59 percent) of mentors matched with youth high on both types of risk reported frequently engaging in these activities. This proportion was significantly higher than that for both groups that were low on individual risk (49 percent).

28. Eight percent of mentors matched with the highest-risk youth reported that their mentee had input on the relationship's focus. This was significantly lower than both groups that were low on environmental risk (14 percent for both).

29. In the study of the BBBSA school-based mentoring program, mentors reported slightly higher levels of closeness (62 percent "agreed" or "strongly agreed" that they felt close to their mentee), and youth reported lower levels of closeness (51 percent reported feeling "very close" to their mentor).

30. Note, however, on average, youth relatively high on both individual and environmental risk (3.34) reported a significantly stronger growth or goal focus in their matches than did youth who were relatively low on both types of risk (3.23).

31. This proportion is higher than that for youth-reported closeness (presented in Table 3.4) because that score reflected a response of 4.0 out of 4.0 whereas this combined measure includes youth who responded with a 3.0 (or higher) out of 4.0 on all three measures.

32. It should be kept in mind that the youth and mentor ratings of relationship quality reported in this chapter are limited to the youth's current (or most recent) mentoring relationship at follow-up. As a result, they may be somewhat more favorable than would be the case if ratings from earlier matches (which may have been less positive given that they had already closed) had been included for youth who had been rematched with a new mentor by the follow-up assessment.

33. Mentors completed their surveys at the end of the 13-month study period or when their relationship with their mentee ended, if that occurred prior to this time. The quotes reported in this chapter were in response to a general question about mentors' experiences in the program: "*Is there anything else you'd like to tell us about your match or your experiences in your mentoring program?*" Their responses to this question are used to illustrate themes we saw in the quantitative data. Not all mentors responded to this question, and those who did may be a distinct subgroup who had relatively more extreme positive or negative experiences. Thus, these quotes cannot be used to discern how common specific experiences were among the mentors.

34. To preserve anonymity, the names of study participants have been changed to pseudonyms throughout the report.

35. Only 0.1 percent had been involved in three matches.

36. Match length does not always reflect the "true" final length of the match—only the length of the match at follow-up. In many cases, as noted in the text, matches were still active and thus ongoing at follow-up.

37. The average number of months of mentoring received was significantly higher when matches were asked to continue at least 18 months (rather than a minimum commitment of 12 months) as part of the program enhancements that were implemented by the two largest programs during the randomized trial and by one of the other programs (9.94 versus 9.29 months).

38. The percentages of first matches lasting different amounts of time reflect information that was gathered about the status of matches after the date of the youth's follow-up assessment. This approach was taken in this set of analyses to ensure that all included matches had the opportunity to last at least 12 months.

39. The percentage of first matches lasting at least one year (53 percent) is somewhat lower than BBBSA's one-year retention rate of 64 percent. This is due, at least in part, to different methods of defining match length. In this study, match length was defined as the number of days that passed between the day the mentor and youth first met each other and the last day that they met together. BBBSA calculates match length considering time after the match's last meeting in which, for example, the case manager continues to work on re-engaging the match.

40. By the time of the follow-up survey, among those who had experienced a match closure, a total of 11 percent of youth high on individual risk only had been rematched. This group was significantly less likely to get rematched than youth who were high on only environmental risk (28 percent), those low on both types of risk (27 percent) and those high on both types of risk (21 percent). A similar pattern was evident when considering rates of rematching for youth who experienced a match closure at any point in the course of the initiative (that is, even after their 13-month follow-up had passed).

41. These are the 915 mentors completing a survey who were the youth's most recent mentor at follow-up and thus the mentor on whom the youth reported in his or her survey.

42. Thirty percent of mentors matched with youth who were high on only individual risk reported that the program initiated the match's closure. This percentage was significantly higher than that for mentors matched with the lowest-risk youth (13 percent).

43. Two fifths of mentors matched with the lowest-risk youth indicated this reason for closure. They were more likely to indicate this closure reason than mentors matched with youth high on only environmental risk (26 percent) and those matched with youth high on only individual risk (25 percent).

44. Only nine percent of mentors matched with the highest-risk youth reported this closure reason. This was a significantly smaller proportion than the 17 percent of mentors matched with the lowest-risk youth and the 25 percent of mentors matched with youth who were high on only individual risk.

45. One fifth of mentors matched with the highest-risk youth reported this closure reason. They were significantly more likely to report this closure reason than mentors matched with youth with the lowest risk and those matched with youth who were high only on environmental risk (both of which were 10 percent).

46. Fourteen percent of mentors matched with the highest-risk youth reported this closure reason compared with only 4 percent of those matched with the lowest-risk youth.

47. Almost half (49 percent) of mentors matched with the highest-risk youth were willing to be matched again, compared with 35 percent of mentors matched with youth with only environmental risk and 34 percent of those matched with youth with only individual risk.

48. A total of 46 percent of mentors matched with the lowest-risk youth and 45 percent of mentors matched with the highest-risk youth reported this challenge. Both groups were more likely to report this challenge than those mentors matched with youth high on only environmental risk (37 percent).

49. About a quarter of mentors (26 percent) were neutral, and 12 percent disagreed.

50. A total of 43 percent of mentors matched with youth who were high on both types of risk reported that they had little in common with their mentee (that is, responded "neutral," "disagree" or "strongly disagree" to the statement, "My mentee and I share similar interests") compared with 34 percent of mentors matched with youth who were high only on environmental risk.

51. These reports did not differ based on youth risk.

52. In all of these cases, mentors matched with youth high on environmental risk (both groups) were significantly more likely than those matched with the lowest-risk youth to report these challenges; all other group comparisons were not significant. For mentee's preparation for meetings: lowest risk = 21 percent; high environmental risk only = 28 percent; highest risk = 29 percent. For getting support from mentee's family: lowest risk = 22 percent; both high environmental risk and highest risk = 30 percent. For mentee's family asking for too much help: lowest risk = 8 percent; high environmental risk = 15 percent; highest risk = 16 percent.

53. Thirty percent of mentors matched with youth who were high on only individual risk and 32 percent of those high on both individual and environmental risk reported managing the youth's behavior as a challenge. These two groups were significantly more likely than mentors matched with the lowest-risk youth (18 percent) to report this challenge.

54. A quarter of mentors matched with the highest-risk youth reported this challenge compared with 16 percent of mentors matched with the lowest-risk youth.

55. A little more than one quarter (27 percent) of mentors matched with the highest-risk youth reported that their expectations were incorrect compared with 20 percent of those matched with youth high on only environmental risk.

56. A total of 35 percent of mentors matched with the lowest-risk youth reported that their expectations around the family's needs were incorrect compared with 42 percent of mentors matched with youth high on only environmental risk.

57. A little over a quarter (27 percent) of mentors matched with the lowest-risk youth reported that their expectations were incorrect about the time commitment required compared with 20 percent of those matched with youth high on only environmental risk.

58. There were no differences on this variable by risk grouping.

59. As in Chapter 3, the findings reported in this chapter are based on data for the youth's current (or most recent) match at the time of the follow-up assessment.

60. Both of the programs that reported that pre-match training was not required did require mentor orientation before matching and presented training as an expectation to their volunteers—but the latter did not have to be completed within one month of the match's start date.

61. Only one program offered online training at the start of the study. Another program began using online training over the course of the study.

62. This proportion is higher when considering mentor reports. Almost three quarters (71 percent) of mentors who completed our survey reported receiving early-match training (that is, training or orientation that did not necessarily meet the criteria outlined in the text) before or within one month of the start of their match. In addition to being based on a more liberal definition of training, the mentors' reports are likely inflated for at least two reasons: (1) some mentors may have been reporting on training they received for a previous match (although, see note 68); and (2) these reports are only for those mentors who completed the mentor survey, which may have been a more invested group than those mentors who did not complete the survey (thus, they may have been more likely to have participated in early-match training).

63. Two hours was the most typical, reported for 47 percent of mentors receiving early-match training.

64. Almost two thirds (66 percent) of mentors matched with youth who were high on only individual risk received early-match training, compared with 57 percent of mentors matched with the lowest-risk youth; 53 percent of mentors matched with the highest-risk youth; and 51 percent of mentors matched with youth who were high on only environmental risk.

65. It is unclear whether this training was "late" pre-match training, occurring after the first month of the match, or training that was meant to be implemented after the start of the match (that is, ongoing training).

66. The percentage of mentors matched with youth high on only environmental risk who received ongoing training (30 percent) was significantly greater than the percentage of mentors

of the lowest-risk youth receiving this type of training (23 percent). The percentages for mentors of youth high only on individual risk (30 percent) and of the highest-risk youth (26 percent) did not differ from the percentages in any of the other risk profile groups.

67. Mentors of the highest-risk youth received about 4.3 hours of ongoing training compared with 3.5 hours for mentors of youth with only environmental risk.

68. It is possible that some mentors did not receive training during the study because they had already been trained previously (prior to the study) in conjunction with mentoring another youth in the program. However, when excluding those mentors who reported that they had previously been matched with a child through the same program, the proportion receiving either early-match or ongoing training remained at 64 percent.

69. Mentors matched with youth who were high on only individual risk received some sort of training significantly more often (74 percent) than did those matched with the lowest-risk youth (65 percent). Among mentors matched with the highest-risk youth and those matched with youth high only on environmental risk, the percentages receiving some sort of training (68 percent in both cases) did not differ significantly from the percentages receiving training in any of the other risk profile groups.

70. Mentors matched with the highest-risk youth were significantly more likely to report learning how to interact with the mentee's family as a training need (64 percent of this group) than were mentors matched with the lowest-risk youth (53 percent).

71. For training on addressing youth's emotional issues, proportions of mentors reporting this training need are as follows: (1) 43 percent of mentors matched with the lowest-risk youth; (2) 50 percent of mentors matched with youth high on only environmental risk; (3) 55 percent of mentors matched with youth high on only individual risk; and (4) 60 percent of mentors matched with the highest-risk youth. The proportions identifying this need in groups (3) and (4) are significantly larger than those doing so in group (1); group (4) is also significantly higher than group (2). For training on addressing youth's social issues, proportions are as follows: (1) 44 percent for mentors matched with the lowest-risk youth; (2) 46 percent for those matched with youth high on only environmental risk; (3) 55 percent for those matched with youth high on only individual risk; and (4) 58 percent for those matched with the highest-risk youth. The proportions identifying this need in groups (3) and (4) are significantly larger than group (1); group (4) is also significantly higher than group (2). For training on addressing youth's behavioral issues: (1) 42 percent for mentors matched with the lowest-risk youth; (2) 43 percent for those matched with youth high on only environmental risk; (3) 51 percent for those matched with youth high on only individual risk; and (4) 55 percent for those matched with the highest-risk youth. Group (4) is significantly larger than groups (1) and (2).

72. A total of 62 percent of mentors reported needing help addressing the mentee's behavioral, social or emotional issues (that is, needing help with one or more of these three issues). Differences across risk profiles for this combined variable were similar to those evident for these training needs when considered separately. A total of 71 percent of mentors matched with the highest-risk youth reported needing help in one or more of the three areas. This percentage was significantly higher than that for mentors matched with youth in the lowest-risk group (55 percent) and for those matched with youth high on only environmental risk (59 percent). Mentors matched with youth high on only individual risk (67 percent) were also significantly more likely than mentors matched with the lowest-risk youth to report having any of these three needs.

73. Mentors matched with the highest-risk youth were significantly more likely to report learning how to work with higher-risk youth as a training need (46 percent of this group) than those matched with youth in the lowest-risk group (38 percent).

74. Mentors matched with the highest-risk youth were more likely to report help with navigating social service systems as a training need (27 percent of this group) than those matched with youth who were high on only environmental risk (19 percent).

75. We also looked at mentor reports of receiving training/support across *all* areas in which they reported needing help. Of those mentors who reported having a need for training/support in one or more of the areas asked about, a little more than half (57 percent) reported having received helpful training in *all* of the areas in which they identified a need.

76. Staff in this program—the university-based program—saw mentors weekly when they met with their mentee, so also communicated with them during these weekly meetings. These interactions, however, were not considered support contacts for purposes of the analyses reported in this chapter.

77. Programs were asked to track every *reciprocated* phone or face-to-face contact with mentors, youth and parents over the course of the youth's study involvement. They also tracked email contacts with the match, but these contacts are not included in the support we discuss in this report.

78. More than a third (36 percent) of mentors matched with youth high on only individual risk reported that their support calls were generally more than 10 minutes. They were significantly more likely to report calls this long than mentors matched with the lowest-risk youth (23 percent) and those matched with youth high on only environmental risk (24 percent).

79. In DuBois et al.'s (2002) meta-analysis, offering a mentor support group was one of the practices in their "theory-based practices index." Higher scores on this index predicted stronger program effects. This practice by itself, however, was not a significant moderator of program benefits in the meta-analysis.

80. About two thirds of mentors matched with youth in the highest-risk category and 66 percent of mentors matched with youth high on only individual risk got support this frequently.

These proportions were both significantly higher than that for mentors matched with youth in the lowest-risk category (58 percent). The proportion for mentors matched with youth in the highest-risk category was also significantly higher than that for mentors matched with youth who were high on only environmental risk (59 percent).

81. For talking with someone at the mentoring program regularly about how things were going with their mentor: The highest-risk youth (82 percent) and those high on only environmental risk (79 percent) were more likely than those high on only individual risk (70 percent) to report that this was "sort of" or "very" true for them. For being able to talk with program staff if they had problems at home or school, those youth high on only environmental risk (78 percent) were significantly more likely than the lowest-risk youth (68 percent) and the highest-risk youth (69 percent) to report that this was "sort of" or "very" true for them.

82. Follow-up surveys were completed only by parents of youth enrolled in the first year of the study at the two largest programs (that is, parents of youth in both the treatment and control groups of the random assignment portion of the study). Responses here are for parents of youth in the treatment group who were matched before follow-up and completed our survey (n = 301).

83. Tests to compare responses across risk groups were not conducted for the parent-reported support variables because the sample sizes for these groups were fairly small.

84. Ideally, we also would have conducted analyses that explored the implications of differences in practices that existed at the level of programs. A key advantage of such analyses is that they avoid the various sources of distortion and bias that can occur when examining practices at the level of individual matches that are discussed in the text (see Wheeler et al. 2010). However, the small number of programs available precluded being able to conduct analyses of practice–outcome associations at the program level.

85. Expectations for frequency of match meetings and match length were also components of the enhancements. But whether these criteria are met may be match- as well as program-driven (and the study did not measure whether and how programs stressed the expectations involved to each match). Thus, they were not included in the "enhancement package" variable. However, because dosage and match length are often considered key match outcomes affected by other practices (for example, more mentor support or better training could encourage more frequent meetings), we did include them as central match outcomes when testing links between match outcomes and other program practices such as mentor training.

86. Matches with youth who were relatively high only on individual risk were significantly more likely to receive at least the partial package of enhancements (45 percent) than matches with youth in each of the other risk profile groups (lowest risk: 34 percent; high only on environmental risk: 30 percent; highest risk: 36 percent).

87. Fifteen percent of those mentors who received early-match training met with their mentee regularly (that is, three or more times during 70 percent or more of the months the match was active) compared with 10 percent of those not receiving early-match training.

88. We found that 15 percent of those mentors who received support calls during 70 percent or more of their months together met this frequently, compared with 9 percent of those who did not receive such consistent support calls.

89. We found that 17 percent of youth who received support calls during 70 percent or more of their months together met this frequently, compared with 11 percent of those who did not receive such consistent support.

90. A total of 15 percent of youth whose parents received support calls during 70 percent or more of their months together met with their mentor frequently (that is, three or more times a month for at least 70 percent of the months that the match was active), compared with 10 percent of those whose parents did not receive such consistent support.

91. We found that 21 percent of matches receiving the full set of program enhancements met this frequently compared with 12 percent not receiving the full set of enhancements. The corresponding percentages were 17 and 10 for the partial set of program enhancements.

92. The analyses examining links between match length and program practices were based on the length of only the youth's first match. The length of these matches was examined past the 13-month follow-up because at 13 months some matches had not had the potential to be together for 12 months (for example, those matches that started even just two months past the youth's baseline survey date).

93. We found that 61 percent of mentors who received early-match training had matches that were 12 or more months long compared with 44 percent of those mentors who did not receive early-match training.

94. Two thirds of mentors who received ongoing training had matches that lasted at least 12 months compared with 48 percent of those mentors who did not receive ongoing training. Note that those mentors in longer matches also had more time to get ongoing training—possibly contributing to this association.

95. A total of 58 percent of mentors who received support calls during 70 percent or more of their months together were in matches that lasted 12 months or longer, compared with 47 percent of those who did not receive such consistent support.

96. We found that 63 percent of matches receiving the partial set of program enhancements lasted at least 12 months, compared with 48 percent of matches not receiving these enhancements.

97. Using an intent-to-treat design, youth were included in the treatment and participant groups whether or not they ultimately received mentoring. See Appendix B for a discussion of this approach and an exception that was made when considering youth in the control group who were served in the second year of the study.

98. Our statistical power to detect differences in program effects across subgroups—in this case, youth with differing risk profiles—was substantially lower than our ability to detect overall program effects. Combining the two groups of mentored youth allowed us to partially offset this limitation.

99. The differential attrition rate (that is, the variation in the proportion of study participants lost to follow-up) across treatment and control groups can represent a serious threat to the integrity of the findings of an evaluation (Flay et al. 2005). However, as noted in the text, rates of attrition across the participant, treatment and comparison groups in the present evaluation were quite similar (see Appendix B for a more detailed discussion of attrition).

100. The outcomes listed constitute the primary outcomes for the evaluation. All of these measures were based on reports by youth. Other secondary outcomes were also examined, including, for the random assignment portion of the evaluation, parent-reported measures of selected outcomes. Appendix B includes an overview of the full set of outcome measures and the results of these additional analyses.

101. We excluded skipping school from this count because only youth who had *not* already skipped school were included in analyses using this measure—our analyses tested which youth *began* to skip school during the study period. Thus, improvement could not be assessed with this measure.

102. As noted in Chapter 1, youth in the comparison group were all enrolled during the first year of the study, whereas those in the participant group were enrolled throughout the study period. In addition, over the course of the study, programs shifted to some extent in their practices, for example, by introducing or expanding their early-match training. In view of these circumstances, we conducted analyses to see whether estimates of program effects differed as a function of the point in time when a youth was enrolled in the study. For the quasi-experimental component of the study, program effect estimates varied significantly for 4 of the 10 primary outcomes: depression, parent trust, misconduct, and number of outcomes showing negative change. For each of these outcomes, program effects appeared to show more favorable effects for mentoring during latter stages of the study. This pattern is consistent with improvements in program practice over the course of the study, but also may reflect other developments taking place over this time frame (for example, improvements in the economy) that also have the potential to influence programs, youth and mentors. Because youth in the comparison group were enrolled only during the first year of the study, the latter changes would disproportionately have influenced youth in the participant group and thus present a potential threat to the accuracy, or internal validity, of our

estimates of program effects. It should be noted, however, that 3 of the 4 outcomes involved were not among those for which a significant program effect was evident in the overall analyses (see Table 5.2), and for the remaining outcome (depression), a significant program effect was nonetheless still evident for youth enrolled during the first year of the study. For the experimental (random assignment) component of the intervention, differences in program effects were evident for 2 of the 10 primary outcomes: prosocial behavior and social acceptance. Although overall program effects were not significant for these outcomes in the random assignment portion of the study, a favorable program effect was evident for social acceptance among those youth in this part of the evaluation who were enrolled earlier in the year during which enrollment took place.

103. Effect sizes were calculated by dividing unstandardized regression coefficients representing the difference between the relevant program participant group (that is, the treatment or quasi-experimental participant group) and the comparison/control group by the standard deviation for the outcome measure (that is, a measure of how great the variation of sample members' scores are around the average score on that measure) at baseline. Illustratively, an effect size of .25 would represent a difference between groups of one fourth of a standard deviation on the outcome measure.

104. The criteria for determining whether an outcome had improved or deteriorated significantly were conservative (see Appendix B for details). Consequently, nearly three quarters (71 percent) of all youth in the full analysis sample for the evaluation (that is, those in the random assignment treatment, quasi-experimental participant and comparison/control groups) did not improve significantly on any of our primary outcome measures; of those who did, most experienced positive change on only one measure (21 percent of the sample). For negative change, 59 percent did not deteriorate significantly in their scores on any measure, and 24 percent worsened on one measure.

105. Effects for both portions of the evaluation remain significant at a .10 level of significance or lower even if one adjusts for having conducted 10 tests using the Benjamini-Hochberg adjustment factor (Benjamini, Hochberg 1995). (See Appendix B for more detail.)

106. These questions were addressed using logistic regressions in which the dependent measures reflected the presence or absence of change on at least one or multiple outcome measures. For instance, if a youth exhibited positive change on one outcome measure, the youth would score positively on the first dependent measure but not the second. A youth who improved on two outcomes would score positively on both measures. Findings indicated that youth in the random assignment treatment group (26 percent) and quasi-experimental participant group (32 percent) were each significantly more likely to show change on at least one outcome measure than were comparison/control group youth (20 percent). The same differences favoring the two mentored groups were

significant when examining the likelihood of improvement on multiple outcomes (3, 7 and 1 percent of youth exhibited this pattern, respectively).

107. We also tested for program benefits in the quasi-experimental portion of the evaluation when excluding from the participant group the 160 youth who also contributed data to the comparison group during their first year of study involvement as members of the control group. Findings did not differ substantively from those reported in Table 5.2, with estimated program benefits reaching statistical significance for the same outcomes in both analyses.

108. An exploratory analysis revealed that youth in both the treatment and participant groups were significantly more likely than youth in the control/comparison group to show significant reductions in their self-reported levels of depressive symptoms, whereas there was no difference between mentored and comparison youth in their rates of worsening over time in their symptoms. It thus appears that the effects observed may have been largely attributable to mentoring *reducing* existing levels of depressive symptoms as opposed to helping youth to avoid symptom onset or worsening.

109. A youth was indicated as showing "serious signs of depression" when he or she had a score of at least 11 on the Short Mood and Feelings Questionnaire used to measure this outcome (SMFQ; Angold et al. 1995).

110. It is also important to note that although research has tended to find associations between self-reported grades and actual grades (for example, Cassady 2001), a recent meta-analysis suggests that this link is not particularly strong, especially for lower-performing youth (Kuncel et al. 2005; see also subgroup analysis in Cassady 2001). The imprecision in our measurement of this outcome arguably would be present equally for our treatment, participant and comparison groups, thus limiting bias in our estimates of program benefits for this outcome. Still, we cannot be sure to what extent the benefits we measured in self-reported grades were reflected in improvements in actual grades.

111. WWC's intervention rating scheme has six mutually exclusive categories that span the spectrum from positive to negative effects: Positive Effects, Potentially Positive Effects, Mixed Effects, No Discernible Effects, Potentially Negative Effects and Negative Effects (*WWC Procedures and Standards Handbook* 2011). The description of our findings according to WWC terminology is based primarily on the WWC criteria for quality of research design (randomized controlled trial versus quasi-experimental) and characterizing study effects; additional considerations, most notably the degree of study attrition and equivalence of intervention and comparison groups at baseline, are also part of the WWC rating framework, but were not systematically applied when arriving at the characterizations that are offered in the text for effects on different outcomes in the current evaluation.

112. These findings, and the pattern evident in the current evaluation, are in line with findings from DuBois et al.'s (2011) meta-analysis in which stronger program effects were evident for programs serving youth with elevated levels of individual, but not environmental, risk relative to those high on both types of risk.

113. For example, in Friends of the Children, a mentoring program that uses paid mentors, all mentors have bachelor's or associate's degrees and previous experience working with youth and are expected to spend several hours with each of their mentees each week.

114. At baseline, youth high on individual risk scored lower on social acceptance, grades and self-perceptions of academic abilities and higher on depression than youth low on individual risk.

115. It is interesting, though, that youth who were high on only individual risk also were the only group to experience a "possible" (albeit not significant) benefit—with an effect size of -.22—on the aggregate negative-change measure. For this reason, a simple "more room to grow" argument seems unlikely to be a primary explanation for the pattern of stronger benefits for this group.

116. Findings from one recent investigation (Schwartz et al. 2011) indicated greater benefits of participation in BBBSA's school-based mentoring program for youth for whom relationship difficulties with parents, peers and teachers were moderate, rather than severe or absent. To explore whether a similar pattern might be evident in our sample, we created a combined measure of the quality of the youth's relationships with parents and peers and tested for both linear and curvilinear moderation of program effects by this measure, the latter being designed to be sensitive to the pattern of results found in the earlier research.

117. We included white, African American and Hispanic youth, omitting other ethnic groups from these analyses because the groups were very small.

118. Potential differences in program effects involving other subgroups, such as youth in foster care and those reporting legal problems (for example, arrest), were of interest but were not investigated due to less than 5 percent of the sample falling into such groups.

119. Despite the potential challenges our findings suggest with regard to mentoring youth with an incarcerated family member, a recent randomized trial evaluation of community-based mentoring for this population did yield some encouraging evidence of program benefits (ICF International 2011).

120. It should be noted that these measures were simple counts of instances of improvement or deterioration that reached a certain absolute threshold, or size. As such, they may have missed more nuanced effects of mentoring, for example, when mentored youth exceeded (or fell below) the thresholds for improvement on outcomes by different amounts than those experienced by youth in the comparison/control group.

121. Random assignment generally provides a less biased estimate of program effects than quasi-experimental designs.

122. The meta-analyses cited were not limited to community-based mentoring programs in which mentors and youth spend time together in varied locations of their choosing. They also included programs that were based in schools or other locations.

123. The greater rate of mentor-youth contact reported in the earlier P/PV study could be attributable, in part, to a reliance on youth reports of contact in that investigation, rather than program records, as was the case in the current evaluation. Yet, in six of the eight sites in the P/PV impact study, programs asked mentors to commit to meeting with their mentee at least weekly for a minimum of three hours at each meeting. These expectations exceed even those for the programs that implemented the targeted enhancements in the current evaluation and thus suggest the potential for a notable difference in actual amounts of mentor-youth contact between the two studies.

124. These findings, and the pattern evident in the current evaluation, are in line with findings from DuBois et al.'s (2011) meta-analysis in which stronger program effects were evident for programs serving youth with elevated levels of individual, but not environmental, risk, relative to those with heightened risk in both individual and environmental areas.

125. This finding again parallels findings of the DuBois et al. (2011) meta-analysis, in which estimated effects were strongest for samples of youth who were high on individual *or* environmental risk, compared with those who were either high or low on both types of risk.

126. As noted, the full "enhancement package" also included different stated expectations about match length and meeting frequency. However, because we did not measure the extent to which each match received these enhancements (that is, whether and how programs stressed these expectations to matches), and these are arguably match outcomes rather than program practices per se, they were not included in our measure of enhancements.

References

Ahrens, K. R., M. Garrison, R. Spencer, D. L. DuBois, P. Lozano and L. P. Richardson
2011   "Qualitative Exploration of the Adult Mentoring Relationships in the Lives of Former Youth in Foster Care." *Children and Youth Services Review*, 33, 1012–1023.

Aiken, L. S. and S. G. West
1991   *Multiple Regression: Testing and Interpreting Interactions*. Newbury Park, CA: Sage.

America's Promise Alliance
2006   *Every Child, Every Promise: Turning Failure into Action*. Washington, DC: America's Promise Alliance. Retrieved 12/20/12 from America's Promise Alliance at www.americaspromise.org/Resources/Partner-Resources/e/Every-Child-Every-Promise.aspx.

Angold, A., E. J. Costello, S. C. Messer, A. Pickles, F. Winder and D. Silver
1995   "Development of a Short Questionnaire for Use in Epidemiological Studies of Depression in Children and Adolescents." *International Journal of Methods in Psychiatric Research*, 5, 237–249.

Angold, A., A. Erkanli, J. Silberg, L. Eaves and E. J. Costello
2002   "Depression Scale Scores in 8–17-year-olds: Effects of Age and Gender." *Journal of Child Psychology and Psychiatry*, 43 (8), 1052–63.

The Annie E. Casey Foundation
2012   *2012 Kids Count Data Book: State Trends in Child Well-Being*. Retrieved 12/13/12 at datacenter.kidscount.org/DataBook/2012/OnlineBooks/KIDSCOUNT2012DataBookFullReport.pdf.

Armsden, G. C. and M. T. Greenberg
1987   "The Inventory of Parent and Peer Attachment: Individual Differences and Their Relationship to Psychological Well-Being in Adolescence." *Journal of Youth and Adolescence*, 16, 427–454.

Baker, D. B. and C. P. Maguire
2005   "Mentoring in Historical Perspective." In D. L. DuBois and M. J. Karcher (eds.). *Handbook of Youth Mentoring*. Thousand Oaks, CA: Sage, 14–29.

Balcazar, F. E. and C. B. Keys
In press   "Goals in Mentoring Relationships." To appear in D. L. DuBois and M. J. Karcher (eds.). *Handbook of Youth Mentoring*. (2nd ed.). Thousand Oaks, CA: Sage.

Barrera, M. and D. D. Bonds
2005   "Mentoring Relationships and Social Support." In D. L. DuBois and M. J. Karcher (eds.). *Handbook of Youth Mentoring*. Thousand Oaks, CA: Sage, 133–142.

Bauldry, S.
2006   *Positive Support: Mentoring and Depression Among High-Risk Youth*. Philadelphia: Public/Private Ventures.

Bauldry, S. and T. A. Hartman
2004   *The Promise and Challenge of Mentoring High-Risk Youth: Findings from the National Faith-Based Initiative*. Philadelphia: Public/Private Ventures.

Benjamini, Y. and Y. Hochberg
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.

Blackwell, L. S., K. H. Trzesniewski and C. S. Dweck
2007   "Implicit Theories of Intelligence Predict Achievement Across an Adolescent Transition: A Longitudinal Study and an Intervention." *Child Development*, 78, 246–263.

Brown, B. B., D. R. Clasen and S. A. Eicher
1986   "Perceptions of Peer Pressure, Peer Conformity Dispositions, and Self-Reported Behavior Among Adolescents." *Developmental Psychology*, 22 (4), 521–530.

Cash, S. J. and J. A. Bridge
2009   "Epidemiology of Youth Suicide and Suicidal Behavior." *Current Opinion in Pediatrics*, 21 (5), 613-619.

Caspi, A., H. Harrington, B. Milne, T. E. Moffit and R. Poulton
2006   "Socially Isolated Children 20 Years Later: Risk for Cardiovascular Disease." *Archives of Pediatrics and Adolescent Medicine*, 160, 805–811.

Cassady, J. C.
2001   "Self-Reported GPA and SAT: A Methodological Note." *Practical Assessment, Research & Evaluation*, 7 (12). Retrieved 12/6/12 from pareonline.net/getvn.asp?v=7&n=12.

Chan, C. S., J. E. Rhodes, S. R. Lowe, S. E. O. Schwartz and C. Herrera
In press   "Pathways of Influence in School-Based Mentoring: The Mediating Role of Parent and Teacher Relationships." *Journal of School Psychology*.

Children's Defense Fund
2012   *The State of America's Children 2012 Handbook*. Retrieved 12/13/12 from Children's Defense Fund at www.childrensdefense.org/child-research-data-publications/data/soac-2012-handbook.html.

Cohen, P. A.
1984   *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Erlbaum.

Cohen, S. and T. A. Wills
1985   "Stress, Social Support, and the Buffering Hypothesis." *Psychological Bulletin*, 98, 310-57.

Cronbach, L. J.
1951   "Coefficient Alpha and the Internal Structure of Tests." *Psychometrika*, 16, 297–334.

Dryfoos, J. G.
1990   *Adolescents at Risk: Prevalence and Prevention.* New York: Oxford University Press, Inc.

DuBois, D. L.
2008   *Youth-Report Measure of Growth/Goal Focus in Youth Mentoring Relationships.* Unpublished measure, University of Illinois at Chicago.

DuBois, D. L., R. D. Felner, S. Brand, R. S. Phillips and A. M. Lease
1996   "Early Adolescent Self-Esteem: A Developmental-Ecological Framework and Assessment Strategy." *Journal of Research on Adolescence*, 6 (4), 543–579.

DuBois, D. L., B. E. Holloway, J. C. Valentine and H. Cooper
2002   "Effectiveness of Mentoring Programs for Youth: A Meta-Analytic Review." *American Journal of Community Psychology*, 30 (2), 157–197.

DuBois, D. L. and M. J. Karcher
In press   "Youth Mentoring in Contemporary Perspective." To appear in D. L. DuBois and M. J. Karcher (eds.). *Handbook of Youth Mentoring.* (2nd ed.). Thousand Oaks, CA: Sage.

DuBois, D. L., N. Portillo, J. E. Rhodes, N. Silverthorn and J. C. Valentine
2011   "How Effective Are Mentoring Programs for Youth? A Systematic Assessment of the Evidence." *Psychological Science in the Public Interest*, 12 (2), 57–91.

Eby, L. T., T. Allen, S. Evans, T. Ng and D. L. DuBois
2008   "Does Mentoring Matter? A Multidisciplinary Meta-Analysis Comparing Mentored and Non-Mentored Individuals." *Journal of Vocational Behavior*, 72, 254–267.

Eddy, J. M., J. Cearley, J. Bergen and J. Stern-Carusone
In press   "Children of Incarcerated Parents." To appear in D. L. DuBois and M. J. Karcher (eds.). *Handbook of Youth Mentoring.* (2nd ed.). Thousand Oaks, CA: Sage.

Elledge, L. C., T. A. Cavell, N. T. Ogle and R. A. Newgent
2010   "School-Based Mentoring as Selective Prevention for Bullied Children: A Preliminary Test." *Journal of Primary Prevention*, 31, 171–187.

Flay, B. R., A. Biglan, R. F. Boruch, F. G. Castro, D. Gottfredson, S. Kellam, E. K. Mościcki, S. Schinke, J. C. Valentine and P. Ji
2005   "Standards of Evidence: Criteria for Efficacy, Effectiveness and Dissemination." *Prevention Science*, 6, 151–175.

Gerard, J. M. and C. Buehler
2004   "Cumulative Environmental Risk and Youth Problem Behavior." *Journal of Marriage and Family*, 66, 702–720.

Good, C., J. Aronson and M. Inzlicht
2003   "Improving Adolescents' Standardized Test Performance: An Intervention to Reduce the Effects of Stereotype Threat." *Journal of Applied Developmental Psychology*, 24 (6), 645–662.

Goodman, R.
1997   "The Strengths and Difficulties Questionnaire: A Research Note." *Journal of Child Psychology and Psychiatry*, 38, 581–586.

Grossman, J. B., C. Chan, S. Schwartz and J. E. Rhodes
2012   "The Test of Time in School-Based Mentoring: The Role of Relationship Duration and Re-Matching on Academic Outcomes." *American Journal of Community Psychology*, 49, 43–53.

Grossman, J. B. and A. Johnson
1999   "Assessing the Effectiveness of Mentoring Programs." In Jean Baldwin Grossman (ed.). *Contemporary Issues in Mentoring.* Philadelphia: Public/Private Ventures, 24–47.

Grossman, J. B. and J. E. Rhodes
2002   "The Test of Time: Predictors and Effects of Duration in Youth Mentoring Programs." *American Journal of Community Psychology*, 30, 199–219.

Hageman, W. J. and W. A. Arrindell
1993   "A Further Refinement of the Reliable Change (RC) Index by Improving the Pre-Post Difference Score: Introducing RCID." *Behaviour Research and Therapy*, 31 (7), 693–700.

Hamilton, S. F. and M. A. Hamilton
1990   *Linking Up: Final Report on a Mentoring Program for Youth.* New York: Cornell University, College of Human Ecology, Department of Human Development & Family Studies.

Harris, J. T. and M. J. Nakkula
1999   *Match Characteristics Questionnaire (MCQ).* Unpublished measure, Harvard Graduate School of Education.

Harter, S.
1985   *Manual for the Self-Perception Profile for Children.* Denver: University of Denver.

Herrera, C., J. B. Grossman, T. J. Kauh, A. F. Feldman and J. McMaken, with L. Z. Jucovy
2007   *Making a Difference in Schools: The Big Brothers Big Sisters School-Based Mentoring Impact Study.* Philadelphia: Public/Private Ventures.

Herrera, C., T. J. Kauh, S. M. Cooney, J. B. Grossman and J. McMaken
2008   *High School Students as Mentors: Findings from the Big Brothers Big Sisters School-Based Mentoring Impact Study.* Philadelphia: Public/Private Ventures.

Herrera, C., C. L. Sipe and W. S. McClanahan, with A. J. A. Arbreton and S. K. Pepper
2000   *Mentoring School-Age Children: Relationship Development in Community-Based and School-Based Programs.* Philadelphia: Public/Private Ventures.

Ho, D. E., K. Imai, G. King and E. A. Stuart
2011   MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *Journal of Statistical Software*, 42 (8), 1–28.

Hu, L. and P. M. Bentler
1999   "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives." *Structural Equation Modeling*, 6 (1), 1–55.

ICF International
2011   *Mentoring Children Affected by Incarceration: An Evaluation of the Texas Amachi Program.* Washington, DC: Office of Juvenile Justice and Delinquency Prevention.

Janosz, M., M. LeBlanc, B. Boulerice and R. E. Tremblay
1997   "Disentangling the Weight of School Dropout Predictors: A Test on Two Longitudinal Samples." *Journal of Youth and Adolescence*, 26 (6), 733-762.

Ji, P., D. L. DuBois and B. R. Flay
2012   *Social-Emotional and Character Development Scale: Development and Initial Validation with Urban Elementary School Students.* Manuscript submitted for publication.

Jucovy, L.
2002   *Measuring the Quality of Mentor-Youth Relationships: A Tool for Mentoring Programs.* Portland, OR: Northwest Regional Educational Laboratory.

Karcher, M. J.
2004   *W. T. Grant-Funded UTSA Study of Mentoring in the Learning Environment (SMILE): Year 1 Results.* San Antonio: University of Texas at San Antonio.

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, 65–77.

Karcher, M. J. and M. J. Nakkula
2010   "Youth Mentoring with a Balanced Focus, Shared Purpose, and Collaborative Interactions." *New Directions for Youth Development*, (126), 13–32.

Keller, T. E.
2005   "The Stages and Development of Mentoring Relationships." In D. L. DuBois and M. J. Karcher (eds.). *Handbook of Youth Mentoring.* Thousand Oaks, CA: Sage, 82–99.

Korenman, S., J. E. Miller and J. E. Sjaastad
1995   "Long-Term Poverty and Child Development in the United States: Results from the NLSY." *Children and Youth Services Review*, 17, 127–155.

Kuncel, N. R., M. Crede and L. L. Thomas
2005   "The Validity of Self-Reported Grade Point Averages, Class Ranks, and Test Scores: A Meta-Analysis and Review of the Literature." *Review of Educational Research*, 75 (1), 63–82.

Lipsey, M. W.
1990   *Design Sensitivity: Statistical Power for Experimental Research.* Newbury Park, CA: Sage.

Malhotra, S. and P. P. Das
2007   "Understanding Childhood Depression." *Indian Journal of Medical Research*, 125, 115–128.

Marsh, H. W., Z. Wen and K. T. Hau
2004   "Structural Equation Models of Latent Interactions: Evaluation of Alternative Estimation Strategies and Indicator Construction." *Psychological Methods*, 9, 275–300.

MENTOR
2006   *Mentoring in America 2005: A Snapshot of the Current State of Mentoring.* Retrieved 4/23/07 from National Mentoring Partnership at http://www.mentoring. org/downloads/mentoring_523.pdf.

Moore, K. A.
2006   *Cumulative Risk Factors Among American Children (Research Brief).* Washington, DC: Child Trends.

Moore, K. A., S. Vandivere and Z. Redd
2006   "A Sociodemographic Risk Index." *Social Indicators Research*, 75, 45–81.

Morrow, K. V. and M. B. Styles
1995   *Building Relationships with Youth in Program Settings: A Study of Big Brothers/Big Sisters.* Philadelphia: Public/Private Ventures.

Muthén, L. K. and B. O. Muthén
1998–2012 Mplus User's Guide. (7th ed.). Los Angeles, CA: Muthén & Muthén.

Muthén, B. O. and A. Satorra
1995 "Complex Sample Data in Structural Equation Modeling." *Sociological Methodology*, 25, 267–316.

Nakkula, M. J. and J. T. Harris
2005 "Assessment of Mentoring Relationships." In D. L. DuBois and M. J. Karcher (eds.). *Handbook of Youth Mentoring*. Thousand Oaks, CA: Sage, 100–117.

Nakkula, M. J. and J. T. Harris
In press "Assessing Mentoring Relationships." To appear in D. L. DuBois and M. J. Karcher (eds.). *Handbook of Youth Mentoring*. (2nd ed.). Thousand Oaks, CA: Sage.

Newcomb, M. D. and P. M. Bentler
1988 "Impact of Adolescent Drug Use and Social Support on Problems of Young Adults: A Longitudinal Study." *Journal of Abnormal Psychology*, 97, 64–75.

Parker, J. G. and S. R. Asher
1987 "Peer Relations and Later Personal Adjustment: Are Low-Accepted Children at Risk?" *Psychological Bulletin*, 102 (3), 357–389.

Parra, G. R., D. L. DuBois, H. A. Neville, A. O. Pugh-Lilly and N. Povinelli
2002 "Mentoring Relationships for Youth: Investigation of a Process-Oriented Model." *Journal of Community Psychology*, 30 (4), 367–388.

The Pew Charitable Trusts
2010 *Collateral Costs: Incarceration's Effect on Economic Mobility*. Washington, DC: The Pew Charitable Trusts.

Posner, J. K. and D. L. Vandell
1994 "Low-Income Children's After-School Care: Are There Beneficial Effects of After-School Programs?" *Child Development*, 65 (2), 440–456.

Rhodes, J. E.
2005 "A Model of Youth Mentoring." In D. L. DuBois and M. J. Karcher (eds.). *Handbook of Youth Mentoring*. Thousand Oaks, CA: Sage, 30–43.

Rhodes, J. E. and D. L. DuBois
2006 "Understanding and Facilitating the Youth Mentoring Movement." *Social Policy Report: Giving Child and Youth Development Knowledge Away*. Retrieved 12/20/12 from Society for Research in Child Development at www.srcd.org/sites/default/files/documents/20-3_youth_mentoring.pdf.

Rhodes, J. E., J. B. Grossman and N. L. Resch
2000 "Agents of Change: Pathways Through Which Mentoring Relationships Influence Adolescents' Academic Adjustment." *Child Development*, 71 (6), 1662–1671.

Rhodes, J. E., R. Reddy, J. Roffman and J. B. Grossman
2005 "Promoting Successful Youth Mentoring Relationships: A Preliminary Screening Questionnaire." *Journal of Primary Prevention*, 26 (2), 147–167.

Rosenbaum, P. and D. Rubin
1985 "Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score." *American Statistician*, 3, 33–38.

Saito, R.
2001 *What's Working? Tools for Evaluating Your Mentoring Program*. Minneapolis: Search Institute.

Sameroff, A. J., W. T. Bartko, A. Baldwin, C. Baldwin and R. Seifer
1998 "Family and Social Influences on the Development of Child Competence." In M. Lewis and C. Feiring (eds.). *Families, Risk, and Competence*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc., 161–185.

Schwartz, S., J. Rhodes, C. Chan and C. Herrera
2011 "The Impact of School-Based Mentoring on Youth with Different Relational Profiles." *Developmental Psychology*, 47, 450–462.

Snyder, C. R., B. Hoza, W. E. Pelham, M. Rapoff, L. Ware, M. Danovsky, L. Highberger, H. Ribinstein and K. J. Stahl
1997 "The Development and Validation of the Children's Hope Scale." *Journal of Pediatric Psychology*, 22 (3), 399–421.

Spencer, R.
2007 "It's Not What I Expected": A Qualitative Study of Youth Mentoring Relationship Failures." *Journal of Adolescent Research*, 22 (4), 331–354.

Tierney, J. P. and J. B. Grossman with N. L. Resch
1995 *Making a Difference: An Impact Study of Big Brothers Big Sisters*. Philadelphia: Public/Private Ventures.

Tolan, P., D. Henry, M. Schoeny and A. Bass
2008 "Mentoring Interventions to Affect Juvenile Delinquency and Associated Problems." *Campbell Systematic Reviews*, 16.

Vandell, D. L.
2003 *Study of Promising After-School Programs Student Survey Glossary*. Retrieved 5/3/07 from Wisconsin Center for Education Research at http://childcare.wceruw.org/pdf/pp/pp_glossary_student_survey_4_5_06.pdf.

What Works Clearinghouse (WWC), US Department of Education's Institute of Education Sciences
2011	*Procedures and Standards Handbook* (Version 2.1). Retrieved 12/12/12 at ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v2_1_standards_handbook.pdf.

Wheeler, M., T. Keller and D. L. DuBois
2010	"Review of Three Recent Randomized Trials of School-Based Mentoring." *Social Policy Report*, 24, 1–21. Retrieved 12/12/12 at www.mentoring.org/downloads/mentoring_1273.pdf.

Wyman, P. A., W. Cross, C. H. Brown, Q. Yu, X. Tu and S. Eberly
2010	"Intervention to Strengthen Emotional Self-Regulation in Children with Emerging Mental Health Problems: Proximal Impact on School Behavior." *Journal of Abnormal Child Psychology*, 38, 707–720.

Appendices

Appendix A: Study Method

This appendix describes the study's design, data collection procedures and the measures used to address study questions. First, we describe the overall design of the study. Then, we discuss the sources of data for the study and the timing of their collection. In the last section, we describe the development of the youth, parent and mentor surveys.

Research Design

This study includes two components: a random assignment impact study and a quasi-experimental study. Each is described below.

Impact study. In the first year of the study, the two largest programs participated in a randomized controlled trial or "impact study." These two programs implemented the practice enhancements described in Chapter 4 for the youth enrolled in this portion of the study.

During the enrollment process, youth who were eligible for program participation (and their parents) were told about the study and the information that would be collected from them over the 13-month study period. Families who agreed to be a part of the study were asked to sign an informed assent form (for youth) and consent form (for parents).[1] After agreeing to participate, both youth and parents completed a baseline survey. Our survey firm, Survey Research Management (SRM), then randomly assigned the youth to one of two groups—a treatment group that was offered the opportunity to receive a mentor as soon as one was available or a control group that was not allowed to receive a mentor through the program until 13 months later.[2] To ensure that youth were matched as soon as possible after completing their baseline, often youth and parents did not complete the baseline survey until a mentor (or potential mentor) had been located. Youth were then sent through random assignment and those who were assigned to the treatment group were matched with a mentor. About two thirds (66 percent) of youth who were selected to receive a mentor were matched within a month of completing their baseline (the average length of time between baseline and matching in the impact study was 35 days).

Parents and youth (both those in the treatment group and those in the control group) were contacted by SRM 13 months later to complete a follow-up survey by phone. At this point, youth in the control group who were still interested in receiving a mentor were given priority to be matched and thus became part of the participant group for the quasi-experimental study (described below).

Random assignment is the best available method for ensuring that the only difference between a treatment and control group is the former group's exposure to the treatment (in this case, access to participation in the mentoring program) and that all other characteristics of the youth will be, on average, the same across the two groups (see Appendix B for a description of this baseline comparison).[3] Comparing the progress made by the treatment and control groups over time allows us to see whether those youth who were offered mentors made more progress than those who were not. At any point in time, the experience of the control group represents what the treatment group would have experienced had they not had the option to enroll in the program.

Quasi-experimental study. Youth in the quasi-experimental study went through a similar enrollment process as those in the impact study; however, following the baseline survey, all youth were included in the "participant group" and were matched as soon as a suitable mentor was found. This approach was taken by five of the seven agencies throughout enrollment for the entire evaluation and by the last two agencies (those participating in the impact study) starting in the second year of the study. Again, agencies tried to ensure that youth were matched soon after baseline and, in many cases, waited to administer the baseline survey until a potential mentor had been found, at which point, the youth's name was sent to SRM and he or she was officially enrolled in the study. More than half (60 percent) of the youth in the quasi-experimental study were matched within 30 days of their baseline assessment, with an average of 47 days between baseline and matching.

Each participating youth was then asked to complete a follow-up survey 13 months after completing the baseline assessment, allowing us to compare their progress over this amount of time with that of youth in the control group from the impact study described above (in this context, this group is referred to as the "comparison group").

After they completed their follow-up assessment, youth in the control group who elected to receive a mentor became members of the participant group in the quasi-experimental study (see details on how our analyses reflected this setup in Appendix B). They then completed a second follow-up survey 13 months later.

General analytic approach. For the main outcome analyses, separate estimates of program effects were derived based on the data collected for the quasi-experimental and impact portions of the evaluation. Doing so allowed us to examine program effects across two samples of youth participants,

using the same control/comparison group. For all other analyses in the report, the two groups of study participants (the treatment group from the impact study and the participant group from the quasi-experimental study) were combined to provide a larger sample to test other study questions, including those about relationship quality, match length, program practices and differences in program experiences and effects as a function of youth risk profile. Thus, in these analyses, some of the members of the combined "program participant" group had gone through random assignment and others had not.

Data Sources

In addition to the youth and parent surveys noted above, the report also drew from a survey of mentors, a program survey, agency records about each match and a staff survey about the case managers working with each match. Table A.1 summarizes these data sources and the timing of their collection. Additional details about each data source are also described below.

Table A.1
Data Sources and Timing

Type of Data	Baseline: Youth's Study Enrollment	Follow-up: 13 Months After Youth Baseline
Youth survey	x	x
Parent survey (impact study)	x	x
Parent survey (quasi-experimental study)	x	
Mentor survey		x (if match closed prior to 13 months, survey was collected then)
Program survey	Completed by agency staff each year of the study to report on agency practices.	
Agency match database	Maintained throughout the study and updated monthly for each match, including, for example, match start and end dates, match meetings, receipt of mentor training and support contacts.	
Case manager survey	Completed at the end of the study by the supervisors of case managers who worked with participating matches.	

Baseline: youth and parent surveys. As noted, each youth in the study completed a baseline survey.[4] This survey was administered one-on-one by agency staff in the program office (or at the youth's home in a quiet, private space) while the youth's parent completed the parent baseline survey in another room. Youth completed the survey online as staff read each question out loud. The survey session took about 30 minutes. A total of 764 youth (both treatments and controls) and their parents completed a baseline survey as part of the impact study, and 777 youth completed a baseline survey for the quasi-experimental study. This latter group included 231 youth who accepted the offer of receiving a mentor following their 13 months in the control group of the impact study.[5] The follow-up survey for these youth (as part of the impact study) also served as their baseline for the quasi-experimental study. Thus, for these youth, the baseline for their second year of study involvement was administered by phone as described below.

Follow-up: youth, parent and mentor surveys. Efforts were made to collect follow-up data from all youth study participants 13 months after their baseline assessment regardless of their match status over the course of the 13-month period. Analyses for both the impact study and the quasi-experimental study were based on all participants for whom we were able to collect follow-up data. At this time, surveys were also administered to parents of youth in the impact study and to mentors (in ongoing matches).

Surveys were administered by the research team with help from participating agencies. All youth in the impact study were contacted by phone by SRM. Youth in the quasi-experimental study were contacted by the programs and surveyed in the office (in a similar setup to that used for the baseline). Agencies were asked to ensure that an unfamiliar staff person administered the follow-up youth surveys (as opposed to the youth's case manager). This procedure was used for the first several months of data collection. However, the agencies had great difficulty reaching youth and arranging for them to come to the office. Thus, the research team began administering the survey by phone to all remaining participants. Rates of success in obtaining follow-up data ranged only slightly across the random assignment treatment group (81 percent), the quasi-experimental participant group (79 percent) and the comparison group (83 percent). All youth in the impact study who completed a follow-up survey did so by phone. In the quasi-experimental study, 180 completed the survey in the agency's office, and 435 completed the survey by phone (mode of survey administration was controlled for in all outcome analyses as described in Appendix B). Youth in the impact study were given a $10 incentive for their completion of the follow-up survey.

The youth and parent surveys included many of the same questions asked at baseline to enable us to assess change over time in key outcomes. Additional questions were also included, asking about experiences in the program and (for youth) in the mentoring relationship, as described in the next section.

Volunteer recruitment for the study began in Spring 2008. Volunteers completed a consent form inviting them to participate in the study. Only volunteers who agreed to participate in the study were matched with participating youth.[6] Thirteen months after the youth's baseline survey, or as the match was closing (for matches that closed prior to 13 months), the mentor was invited to complete the mentor survey—either online or on paper (almost all completed it online). If the mentor was still meeting with the youth or in touch with the agency, then agency staff contacted the mentor to remind him or her to complete the survey; if not, the research team sent the mentor an email invitation to do so. Mentors whose matches were closing were asked to complete the survey immediately after their matches closed (to ensure that we did not lose contact with these mentors and that they could remember the details of their match interactions). Mentors were given a $10 incentive for their completion of the follow-up survey.

The survey asked about the mentor's background (for example, demographics, volunteer experience, marital and student status); past experience interacting with and mentoring youth; goals for the mentoring relationship; and opinions about youth. We also asked mentors about the level and types of training and support they received through their participation; the types of activities they engaged in with their mentee; and the quality of their relationship. If the match had ended (or was ending), we asked mentors why the match was ending and how they felt about this.

In some instances, by the time of the follow-up, youth had experienced a match closure and had been rematched with a new mentor. For the most part, study analyses were limited to data obtained from the mentor who was matched with the youth at the time of follow-up or, if there was no active match, the mentor who had been matched most recently with the youth. Youth were asked to report on their relationship with their current (or most recent) mentor in their follow-up survey. Limiting consideration to survey data obtained from this mentor thus ensured that findings would refer to the same relationship that was reported on by youth. Of a total of 904 mentors that fit this description, 85 percent completed the follow-up survey.

Program survey. Staff from each of the seven participating agencies completed a survey asking about program structure, recruitment, training, supervision and other program characteristics. Data from this survey were used in Chapter 4 and Appendix C to describe program characteristics (for example, methods for recruitment, matching and training). Program surveys were self-administered at three time points: (1) at the beginning of the study (in 2008), asking about agency practices prior to the study's beginning (baseline); (2) in 2010, asking about practices in 2009; and (3) in 2011, asking about practices in 2010. Research liaisons (that is, one staff member at each agency who helped oversee evaluation activities at the agency) completed the survey, but were asked to recruit agency staff to help complete sections with which they were unfamiliar.

Match database. Throughout the study, each agency kept an Excel database to track information about participating youth and matches. This information included the date and length of each mentor's early-match training, match start and end dates (that is, the first and last days that the match met) and mentor attendance at ongoing training. The database was also updated monthly to indicate for each match whether the mentor, youth and parent had been reached by case managers for support contacts and, if so, how many contacts had been made and through what means (that is, phone or face-to-face versus email). These monthly updates also included information about how many times the match had met, how long each meeting lasted, and whether the match had any non-face-to-face meetings (for example, email, phone) and how many. The data from these databases are used to summarize the frequency and duration of match meetings in Chapter 3 and the mentor training and support provided to matches in Chapter 4.

Case manager survey. Staff supervisors were asked to provide information about each case manager who provided support to participating matches—the case manager's dates of work with each match as well as information about the case manager's background (for example, tenure at the agency, educational background) and more general perceptions of the quality of her or his work across all participating matches. These data and the analyses to which they contributed are discussed in more detail in Appendix G.

Youth, Parent and Mentor Survey Development

The youth, parent and mentor surveys were developed with four overarching goals in mind: (1) to measure youth risk with as much detail as possible (these questions were included mainly in the parent survey); (2) to assess participants' experience of program practices (such as training and support); (3) to explore the experiences of mentors and youth within their mentoring relationships; and (4) to assess effects of program participation on youth outcomes.

For the first goal, we turned to the existing literature on risk to ensure that our questions encompassed a wide range of risk factors and were relevant for this particular population and age group of youth. We wanted to make the risk assessment as brief and easy to complete as possible so that

programs could use it to assess risk efficiently at intake (see Appendix D for more details on how risk was assessed). For goals two and three, we relied on several existing measures of relationship quality and perceptions of program training and support. We also developed several new questions, including those about the mentor's background, perceived challenges, training and support needs, and mentor perceptions of the youth's risk level. The purpose of these latter questions was to allow an assessment of how risk may have played a role in the developing relationship and in the mentor's needs for support. Finally, for the fourth goal, we developed the outcome measures for the study considering: (1) existing theories of how mentoring works (for example, Rhodes 2005); (2) the outcomes that the literature suggests are most likely to be affected by mentoring; and (3) those outcomes of greatest import to funders and policymakers. We reviewed a number of survey instruments that measured these outcomes, selecting scales that were appropriate for use with youth from ages 9 to 15 and their parents and, where possible, had been used in previous mentoring evaluations to allow comparison across studies. When suitable measures were not found, we developed our own.

Combining these measures, we developed drafts of the youth and parent surveys, which included all outcome, risk and background measures as well as, at follow-up, measures of program support and, in the youth survey, relationship quality. We also developed a mentor survey with information on the mentor's background, match activities, match support, training and relationship development.

All measures and the survey instruments that resulted from them were selected and developed by the P/PV researchers in close consultation with Dr. David DuBois and Dr. Janet Heubach from Washington State Mentors. In addition, the study's procedures, all surveys, consent forms and youth assent forms were reviewed, revised and approved by an independent, federally registered Institutional Review Board.

Selected study measures. All selected study measures are single items or scales that have been validated in prior studies and/or used in previous P/PV evaluations or created specifically for this study. Appendix D describes the items used to assess youth risk. Tables A.2 and A.3 (at the end of this appendix) describe key relationship quality and support measures included in the youth and mentor follow-up surveys.

Table A.2 also describes the outcome measures used in the study. The baseline and follow-up surveys included 19 youth-reported and 3 parent-reported outcomes falling into 8 domains. The analyses presented in the main text of the report use only a subset of these measures designated as "primary" (one from each of the 8 domains). These eight outcome measures (italicized in Table A.2) form the basis for the outcome analyses reported in Chapter 5. (See Appendix B for a detailed description of the full set of outcome measures and how we selected our primary outcomes.)

Reliability. When appropriate, the reliability, or internal consistency, of each outcome measure was assessed for the study sample at both baseline and follow-up. The reliability of a scale refers to how consistently the items measure an underlying construct (for example, self-esteem, social acceptance). Coefficient alpha (Cronbach 1951) is a statistic used to assess this "internal reliability." Alpha values range from 0 (indicating no internal consistency: The items have nothing in common) to 1 (indicating perfect consistency among the items). We consider values above .70 to be acceptable. For some outcome measures, it was not appropriate or possible to compute a reliability estimate using coefficient alpha either because the measure consisted of only one item or was scored in a manner that did not involve a simple averaging of responses across the different items. Alpha values for the remaining outcome measures all reached or approached acceptable levels, ranging from .66 to .88 at baseline and .67 to .89 at follow-up (see Tables A.2 and A.3).

Table A.2

Measure Descriptions and Reliability for Youth- and Parent-Reported Outcomes and Relationship-Quality Measures

Constructs	Title of Measure	Author(s) of Measure	Response Stem and Sample Items	Response Choices	Number of Items	Scoring	Cronbach's Alpha (α)[a]
OUTCOMES							
Psychological well-being							
Depression	Short Mood and Feelings Questionnaire (SMFQ)	Angold et al. (1995)	In the past two weeks… • I felt miserable or unhappy. • I thought nobody really loved me.	1 = Not true to 3 = True most of the time	13	Mean of items	.88/.89
Self-worth	Self-Esteem Questionnaire	DuBois et al. (1996)	• I like being just the way I am. • I am happy with myself as a person.	1 = Not at all true to 4 = Very true	8	Mean of items	.74/.77
Hope	Children's Hope Scale	Snyder et al. (1997)	• I am doing just as well as other kids my age. • When I have a problem, I can come up with lots of ways to solve it.	1 = None of the time to 6 = All of the time	6	Mean of items	.81/.78
Emotional symptoms[b]	From Strengths and Difficulties Questionnaire (SDQ)	Goodman (1997)	Over the last six months how true is it that your child… • Is often unhappy, downhearted or tearful. • Has many worries, often seems worried.	1 = Not true to 3 = Certainly true	5	Mean of items	.75/.71
Prosocial behavior and activities							
Prosocial behavior	Social-Emotional and Character Development Scale	Ji et al. (2012)	• I treat my friends the way I like to be treated. • I am nice to kids who are different from me.	1 = Not at all true to 4 = Very true	6	Mean of items	.74/.68
Prosocial behavior[b]	From the SDQ	Goodman (1997)	Over the last six months how true is it that your child… • Is considerate of other people's feelings. • Is helpful if someone is hurt, upset or feeling ill.	1 = Not true to 3 = Certainly true	5	Mean of items	.71/.67
Honesty	Social-Emotional and Character Development Scale	Ji et al. (2012)	• I tell others the truth. • I admit my mistakes.	1 = Not at all true to 4 = Very true	5	Mean of items	.68/.68
Community service		Public/Private Ventures	How often in the past three months have you… • Done some type of community service or volunteer work?	0 = Never in my life to 3 = I did it 3 or more times in the last 3 months	1	NA	
Peer relationships							
Social acceptance	Adapted from the Self-Perception Profile for Children	Harter (1985)	• I have a lot of friends. • I wish that more people my age liked me. (reversed)	1 = Not at all true to 4 = Very true	6	Mean of items	.76/.79

Table A.2, *continued*

Measure Descriptions and Reliability for Youth- and Parent-Reported Outcomes and Relationship-Quality Measures

Constructs	Title of Measure	Author(s) of Measure	Response Stem and Sample Items	Response Choices	Number of Items	Scoring	Cronbach's Alpha (α)[a]
Adult relationships							
Parent trust	Adapted from the Inventory of Parent and Peer Attachment	Armsden, Greenberg (1987)	• I trust my parent. • My parent accepts me as I am.	1 = Hardly ever to 4 = Pretty often	7	Mean of items	.84/.85
Non-parental help with college/jobs	Developed for this study		How often in the past three months have you… • Talked with an adult (not your parent) about what you need to do to get a good job?	0 = Never in my life to 3 = I did it 3 or more times in the last 3 months[c]	3	Mean of items with 0 and 1 combined (scored as 1)	.74/.73
Number of non-parental special adults		Adapted from Herrera et al. (2007)	• Right now in your life, is there a special adult (not your parent or guardian or other person who has raised you) who you often spend time with? A special adult is someone who really cares about what happens to you. It is someone (a) who you look up to and encourages you to do your best, (b) who influences what you do and the choices you make, and (c) who you can talk to about personal problems.	0 = No 1 = Yes	1	If no = 0; if yes = number of choices endorsed from 8 options (e.g., neighbor, teacher), after excluding "an adult mentor you are matched with through a program"	NA
Problem behavior							
Misconduct		Brown et al. (1986); adapted by Posner, Vandell (1994)	How often in the past three months have you… • Taken something on purpose that didn't belong to you? • Hit someone because you didn't like something they said or did?	0 = Never in my life to 3 = I did it 3 or more times in the last 3 months[c]	5	1 for response of "2" or "3" for any of the questions; 0 otherwise	NA
Substance use		Policy Studies Associates	How often in the past three months have you… • Drunk alcohol without your parents knowing?	0 = Never in my life to 3 = I did it 3 or more times in the last 3 months[c]	5	1 for response of "1", "2" or "3" for any of the questions; 0 otherwise[d]	NA
Legal problems	Developed for this study		• How often in the past three months have you been sent away or placed in juvenile home by the court?	0 = Never in my life to 3 = I did it 3 or more times in the last 3 months[c]	2	1 for response of "1", "2" or "3" for either question; 0 otherwise[d]	NA
Conduct problems[b]	From the SDQ	Goodman (1997)	Over the last six months how true is it that your child… • Often lies or cheats. • Often fights with other children or bullies them.	1 = Not true to 3 = Certainly true	5	Mean of items	.72/.74

Table A.2, *continued*

Measure Descriptions and Reliability for Youth- and Parent-Reported Outcomes and Relationship-Quality Measures

Constructs	Title of Measure	Author(s) of Measure	Response Stem and Sample Items	Response Choices	Number of Items	Scoring	Cronbach's Alpha (α)[a]
Academic attitudes							
Self-perceptions of academic abilities	Adapted from the Self-Perception Profile for Children	Harter (1985)	• I often forget what I learn. (reversed) • I am very good at my schoolwork.	1 = Not at all true to 4 = Very true	6	Mean of items	.66/.73
School liking	Adapted from a scale tested with middle-school youth in a project conducted by Jacquelynne Eccles	Jacquelynne Eccles	• In general, I like school a lot. • I look forward to going to school every day.	1 = Not at all true to 4 = Very true	3	Mean of items	.83/.84
Educational expectations		Adapted from Vandell (2003)	How sure are you that you will… • Finish high school? • Go to college or some other kind of school beyond high school?	1 = Not at all sure to 4 = Very sure	2	Mean of items	.74/.71
Academic behaviors							
Hours worked on homework	Developed for this study		• This year, on an average school day, how many hours after school have you worked on homework or projects for school?	0 = I do not do any homework or school project work on an average school day to 4 = 3 or more hours per day	1	NA	
Skipping school		Adapted from Herrera et al. (2007)	How often in the last three months of school have you… • Skipped school without your parent or other person who has raised you knowing? • Skipped a class without being allowed?	0 = Never in my life to 3 = I did it 3 or more times in the last 3 months of school[e]	2	1 for response of "1", "2" or "3" for either question; 0 otherwise[d]	NA
Academic performance							
Grades		Developed for BBBSA's Youth Outcomes Survey by Public/Private Ventures and Jean Rhodes	• Thinking about the grades and marks you are getting in school, please circle how you are doing in each of the following areas (i.e., math, reading or language arts, social studies, science).	1 = F (not very good at all) to 5 = A (excellent)	4	Mean of items	.78/.78

Table A.2, *continued*

Measure Descriptions and Reliability for Youth- and Parent-Reported Outcomes and Relationship-Quality Measures

Constructs	Title of Measure	Author(s) of Measure	Response Stem and Sample Items	Response Choices	Number of Items	Scoring	Cronbach's Alpha (α)[a]
			RELATIONSHIP QUALITY				
Closeness to mentor		Herrera et al. (2007)	• How close do you feel to your mentor?	1 = Not close at all to 4 = Very close	1	NA	
Growth/goal focus		DuBois (2008)	• My mentor and I talk together about how to solve problems. • My mentor and I work on projects together.	1 = Not at all true to 4 = Very true	6	Mean of items	.83
Youth centered		Adapted from Grossman, Johnson (1999)	• My mentor and I do things I really want to do. • My mentor thinks of fun and interesting things to do.	1 = Not at all true to 4 = Very true	5	Mean of items	.71

Note: Primary outcome measures are italicized in the table.

[a] In cases where the measure was collected at both baseline and follow-up, the alpha for the baseline assessment is listed first, followed by the alpha for the follow-up assessment. For all mentor- and youth-reported measures of relationship quality, only the follow-up alpha is listed because these measures were not administered at baseline.

[b] This is a parent-reported measure.

[c] Full response set is as follows: 0 = "Never in my life"; 1 = "I have done this, but not in the last 3 months"; 2 = "I did it 1–2 times in the last 3 months"; 3 = "I did it 3 or more times in the last 3 months."

[d] Analyses of program effects for this outcome focused on the proportion of youth who had initiated the behavior at follow-up and thus were limited to those youth who did not report already having engaged in the behavior at baseline (that is, responded "Never in my life" to each of the relevant questions). The scoring information provided for the measure in the table refers to how it was scored at follow-up in order to determine whether youth reported having engaged in the behavior in the past three months at this time point.

[e] Full response set is as follows: 0 = "Never in my life"; 1 = "I have done this, but not in the last 3 months of school"; 2 = "I did it 1–2 times in the last 3 months of school"; 3 = "I did it 3 or more times in the last 3 months of school."

Table A.3

Measure Information and Reliability for Selected Mentor-Reported Measures

Constructs	Title of Measure	Author(s) of Measure	Response Stem and Sample Items	Response Choices	Number of Items	Scoring	Cronbach's Alpha (α)
Closeness to mentee	From the Match Characteristics Questionnaire (MCQ)	Harris, Nakkula (1999)	• I feel close with my mentee.	1 = Strongly disagree to 5 = Strongly agree	1	NA	
Match activities	Adapted from Match Engagement in Activities Measure	Karcher (2004)	About how much of your time with your mentee did you spend engaging in the following activities… • Fun: Having times when you do nothing but fun things with your mentee? • Talking: Talking about your mentee's family? • Character/behavior change: Teaching your mentee social skills? • Enrichment: Going to cultural events? • Structured/program activities: Participating in program-wide events organized by your mentoring program?	0 = None to 4 = Most	26 total items: Fun (4); Talking (7); Character/ behavior change (4); Enrichment (5); Structured/ program activities (6)	Mean of items for each subscale	Fun: .72 Talking: .85 Character/ behavior change: .84 Enrichment: .66 Structured/ program activities: 78
Staff support		Karcher (2004)	• Program staff have given suggestions on what I can do with my mentee. • Program staff have shared important information with me about my mentee.	1 = Strongly disagree to 5 = Strongly agree	4	Mean of items	.79

Appendix B: Analysis of Program Outcomes

In this appendix we describe our general analytic approach to estimating program effects on youth outcomes and provide details for several additional sets of related analyses referenced in the report.

Analyses of Program Effects on Primary and Secondary Outcomes

The theory of mentoring, as well as prior research, suggests that mentoring can affect a young person in many domains. For this evaluation we considered the effect of mentoring program participation in eight areas:

- Emotional/psychological well-being;

- Prosocial behavior and activities;

- Peer relationships;

- Adult relationships;

- Antisocial behaviors;

- Academic attitudes;

- Academic behaviors; and

- Academic performance.

To assess changes in outcomes in these areas, youth completed surveys both during the application process (baseline) and 13 months later, as described in Appendix A. In these surveys (and a survey of parents), we collected information on 22 outcomes—4 related to emotional/psychological well-being (self-worth, hope, depression and parent reports of emotional symptoms), 4 on prosocial behavior and activities (honesty, prosocial behavior, community service and parent reports of prosocial behavior), 1 on peer relationships (social acceptance), 3 on adult relationships (parent trust, number of non-parental special adults and non-parental help with college/jobs), 4 on antisocial behaviors (substance use, misconduct, legal problems and parent reports of conduct problems), 3 on academic attitudes (self-perceptions of academic abilities, school liking and educational expectations), 2 on academic behaviors (hours worked on homework and skipping school) and 1 on academic performance (self-reported grades).

However, the more outcomes one examines, the more likely it is that a few of them will yield a statistically significant finding—by chance—even though the intervention (in our case, mentoring program participation) did not actually affect youth (or, more technically, would not be found

to affect them if the same study were repeated many times with different samples from the same population of interest). Therefore, without knowing which measures yielded evidence of program effects, we designated one measure in each of our eight broad outcome areas as the primary outcome measure in that domain. We based this selection on the measure's policy relevance and how well it represented each domain. If we found evidence that mentoring affected a primary outcome in the domains we examined, we concluded that mentoring did show evidence of having an effect in that domain; if we did not, we concluded that we do not have evidence of an effect in that domain. Below are the outcome measures (all youth-reported) that we tested in our primary analysis—namely, the analyses upon which we judged the effectiveness of the programs in this evaluation:

- Depression (that is, depressive symptoms—a measure of emotional/psychological well-being);

- Prosocial behavior and activities (prosocial behavior);

- Social acceptance (peer relationships);

- Parent trust (adult relationships);

- Misconduct (antisocial behaviors);[7]

- Self-perceptions of academic abilities (academic attitudes);

- (Initiation of) skipping school (academic behaviors);[8] and

- Grades (academic performance).

The effects of mentoring on the primary outcomes are reported in the main text, while the effects on the other variables are reported in this appendix and are part of our secondary analysis.

In addition to the eight primary outcomes listed above we added two additional primary outcomes—the number of positive (that is, favorable) "reliable changes" experienced by an individual youth over the 13-month study period and the number of negative (that is, unfavorable) "reliable changes" experienced by a youth over the same period. A reliable change measure is an indicator that equals 1 if the observed change is greater than what would be expected due to measurement error (Hageman, Arrindell 1993).[9] For seven of the eight primary outcomes (we excluded skipping school),[10] we examined whether the change from baseline to follow-up fell outside the 90-percent confidence interval. If the change was greater than the upper end of the confidence interval, it was counted as a "reliable," positive

pre-/post- movement, whereas if the change was less than the lower end of the confidence interval, it was counted as a "reliable," negative pre-/post- movement. These additive measures are *not* the number of positive or negative *impacts* a youth experienced (that is, the extent to which the youth changed on a given measure *relative* to the change experienced by youth in a comparison group), but rather are the number of the primary outcomes that appear to have increased or decreased beyond a specified threshold over the 13 months for a given youth. They indicate the number of outcomes for which statistically "reliable" improvement or deterioration is evident for each youth.

Method

To examine the effect of mentoring on youth program participants, we conducted two types of intent-to-treat comparisons:

- Experimental/Random Assignment (RA)—comparing the group of youth who were randomly selected not to receive a mentor (the control group) with the group of youth who were randomly selected to receive a mentor (the treatment group). Both sets of youth applied to the two largest programs in the first year of the evaluation.

- Quasi-Experimental (QE)—comparing the non-mentored control group created experimentally in the first year of the evaluation in two programs (called the "comparison group" in the QE setting because it was not drawn randomly from the exact same population as the participant group) with all eligible youth who applied to any of the seven study programs and were not required at that time to go through the random assignment process (called the "participant group" to distinguish it from the randomly assigned treatment group).

It should be noted that although all youth in the participant and treatment groups joined the pool of potential mentees, not all were ultimately matched.[11]

In the RA portion of the evaluation, 764 eligible youth applied to the program with 379 of them randomly assigned to the treatment group to receive mentors and 385 randomly assigned to the waitlist for 13 months (control group). Analyses for this portion of the evaluation were based on 308 treatment group and 321 control group youth for whom we were able to collect data at the 13-month follow-up. For the QE portion of the evaluation, at the five smaller programs and during the second year in the two larger programs, 777 eligible youth applied to the programs. Case managers tried to match all of these youth, and they comprised the

participant group. Analyses for the QE portion of the evaluation were based on the 615 of these youth for whom follow-up data were obtained and the 321 comparison/control group youth for whom follow-up data were obtained.

Because control group youth could reapply to the program after they completed their 13-month follow-up survey, a substantial number, 231, were included as youth in the participant group. We were able to follow up with 160 of these youth; thus they are in the QE analysis sample twice—once as controls in Year 1 and once as participants in Year 2. If a control group youth did not elect to receive mentoring through the program the second year, the youth is in the dataset only once, as a control. We discuss how we handle this analytically below.

As noted, analyses for both the RA and QE components of the evaluation generally followed an intent-to-treat design, whereby once youth were determined to be eligible for participation, had received parental consent to participate and had completed a baseline survey, they were included as study participants in analyses regardless of whether they ended up receiving mentoring services. This ensures that the follow-up sample is representative of the full sample and not biased by excluding nonrandom groups of youth (for example, those youth in the treatment or participant groups who were not ultimately served or those who had very short matches). There was one exception to this strategy, which involved the 231 youth in the control group from the RA portion of the study who were ultimately matched in their second year of study involvement (and thus are included in the QE participant group's data). We conducted QE follow-up surveys only with these control group youth who wanted a mentor after their first year as controls—not the entire control group to whom this offer was made (as would be expected using an intent-to-treat design). We did this because many of the control group youth no longer wanted a mentor by the time of follow-up, so their inclusion in the participant group would not be consistent with how we had determined inclusion in the overall study. Analyses of program effects were conducted with and without these youth, and we found no changes in the outcomes for which significant effects were indicated in our main analyses (described below).

Analytical Strategy

The intent-to-treat analysis attempts to estimate the average effect that *offering* youth the opportunity of program involvement had on the outcomes described above. For each outcome, the basic model used was:

$$y_{ij} = \beta_0 + \beta_1 \text{Pre}_{ij} + \beta_2 P_{ij} + \beta_3 T_{ij} + \boldsymbol{\beta}_k \mathbf{X}_{ijk} + \sum_j \gamma_j S_j + \varepsilon_{ij} \ (1)$$

for   i = 1, …, N individuals
         j = 1, …, 6 programs
         k = 1, …, K baseline individual-level covariates

where y_{ij} is the outcome of interest for youth i in program j;

         Pre_{ij} is the baseline measure taken at application;

         P_{ij} is an indicator variable equal to 1 if an eligible youth applied to the program when random assignment was *not* being conducted;

         T_{ij} is an indicator variable equal to 1 if youth i in program j was randomly assigned to the treatment group or 0 if the youth was randomly assigned not to receive a mentor for 13 months (that is, assigned to the control group);

         β_w and γ_v are estimated parameters on a covariate (estimates of particular interest are described below);

         β_2 is the "fixed effect" estimate of the program effect from the quasi-experimental data;

         β_3 is the "fixed effect" estimate of the program effect from the experimental data;

         γ_j is the estimated program-specific effect;

         $\mathbf{X}_{ijk}$ is a vector of k baseline youth-level covariates;

         S_j are the program dummies; and

         ε_{ij} is the individual-level error component.

Given that some youth appear twice in the dataset (that is, those youth who were in the control group in their first year of study involvement and part of the participant group in their second year of study involvement), the Year 1 and Year 2 observations are correlated for these individuals. Therefore, we use a maximum likelihood technique to simultaneously estimate the coefficients of equation (1) and the covariance matrix, assuming an AR(1) error process for individuals who are in the dataset twice.

In cases where the outcome is measured dichotomously, we used logistic regression analyses within the covariance framework described above. The dependent variable, y_{ij}, takes the form of the log odds of observing the outcome:

$$y_{ij} = \log\left(\frac{\Phi_{ij}}{1 - \Phi_{ij}}\right) \qquad (2)$$

where Φ_{ij} is the probability of observing the outcome, and 1- Φ_{ij} is the probability of not observing the outcome.

In cases where the outcome is a count, such as the number of outcomes that "reliably" deteriorated over time, we used negative binomial regression analyses within the covariance framework described above. The dependent variable, $\log(y_{ij})$, is assumed to be related to the covariates linearly, and the error term is distributed as a Poisson-Gamma (or negative binomial) distribution:

$$\log(y_{ij}) = \beta_0 + \beta_1 Pre_{ij} + \beta_2 P_{ij} + \beta_3 T_{ij} + \boldsymbol{\beta}_k \mathbf{X}_{ijk} + \sum_j \gamma_j S_j + \varepsilon_{ij} \quad (3)$$

Maximum likelihood techniques are used to estimate the coefficients.

Subgroup Analyses

We also tested whether the mentoring programs had different effects on different types of youth based on their baseline characteristics (for example, their risk profile, gender, age). To increase our power to detect subgroup differences, the treatment group from the randomized portion of the evaluation and the participant group from the quasi-experimental portion of the evaluation were combined in all of our subgroup analyses (this combined group of youth are referred to as "program participants").

To allow the effect of mentoring to vary by the risk subgroups, we replaced the simple participation variables in equations (1) through (3) with three participation dummies reflecting the youth's participation status and his or her environmental and individual risk: (1) a dummy variable that is 1 if the youth was a program participant with high environmental risk; (2) a dummy variable that is 1 if the youth was a program participant with high individual risk; and (3) a dummy variable that is 1 if the youth was a program participant high on both environmental and individual risk. In addition, the model included the general risk profile dummies to ensure that the outcomes of program participants in a particular risk group were being compared with outcomes only of comparison youth with a similar risk profile. Algebraically, the model was:

$$y_{ij} = \beta_0 + \beta_1 Pre_{ij} + \beta_{2E} HIENVT_{ij} + \beta_{2I} HIINDIV_{ij}$$

$$+ \beta_{2EI} ENVT * INDIV_{ij} + \beta_{3E} P * HIENVT_{ij} + \beta_{3I} P * HIINDIV_{ij}$$

$$+ \beta_{3EI} P * ENVT * INDIV_{ij} + \boldsymbol{\beta}_k \mathbf{X}_{ijk} + \sum_j \gamma_j S_j + \varepsilon_{ij}$$

where $HIENVT_{ij}$ is 1 if the youth is high on environmental risk;

         $HIINDIV_{ij}$ is 1 if the youth is high on individual risk;

         $ENVT * INDIV_{IJ}$ is 1 if the youth is high on both environmental and individual risk;

         P is 1 if the youth is a program participant (that is, eligible to be matched in either the randomized or

the quasi-experimental portion of the evaluation) or 0, if the youth is not a program participant;

$P * HIENVT_{ij}$ is 1 if the youth is a program participant with high environmental risk;

$P * HIINDIV_{ij}$ is 1 if the youth is a program participant with high individual risk;

$P * ENVT * INDIV_{ij}$ is 1 if the youth is a program participant with high environmental and individual risk;

β_{2r} (where r is E, I or EI) is the correlation between the outcome and being a youth who is high on environmental risk (E), or being a youth who is high on individual risk (I), or being a youth who is high on both risks (EI), holding the other covariates constant; and

β_{3r} (where r is E, I or EI) is the estimated effect of mentoring on the outcome if the youth is high on environmental risk (E), if the youth is high on individual risk (I), or if the youth is high on both risks (EI), holding the other covariates constant.

If the coefficient, β_{3EI}, in the above model, which represents non-additive differences in program effects in association with youth individual and environmental risk status, was not significant, the model was re-estimated with this term removed to get a better test of additive differences in program estimates associated with whether youth were high on individual, β_{3I}, or environmental risk, β_{3E}.

Most of the other youth characteristics examined for differences in estimated program effects have only two categories (girl/boy, being from a single-parent family or not, etc.). For these other subgroup analyses, we included interaction terms in equation (1) between program participant status and individual-level covariates:

$$y_{ij} = \beta_0 + \beta_1 Pre_{ij} + \beta_2 P_{ij} + \delta_k PX_{ijk} + \beta_k X_{ijk} + \sum_j \gamma_j S_j + \varepsilon_{ij} \quad (4)$$

where PX_{ijk} is the interaction of the participation dummy (P) with one of the two possible subgroup dummy variables, such as being a boy;

β_2 is the estimate of the program's effect that affects both subgroups;

δ_k is the estimate of the differential effect on program participants that fall into the k subgroup; and

X_{ijk} is 1 if the youth has this characteristic, such as being a boy, or being Hispanic. This X is one of k characteristics that was included in the basic regression, $\mathbf{X}_{ijk}$.

Thus, for example, if the interaction is between participation status and being male, then the impact on female participants would be β_2, while the impact for male participants would be $\beta_2 + \delta_f$.

For the youth background characteristics and experiences that were continuous, such as age, X_{ijk}, the subgroup dummy variable in the above equation, was replaced by the continuous variable. When the interaction involving a continuous variable was significant, we followed the Aiken and West (1991) approach of evaluating strength and significance of the estimated program effect at low (1 standard deviation below the sample mean), average and high (1 standard deviation above the sample mean) values of the continuous variable.

Equivalence of the Participant/Treatment and Comparison/Control Groups and Model Covariates

The experimental and quasi-experimental comparisons outlined above yield unbiased estimates of the effect of mentoring as long as the treatment and control or the participant and comparison groups are the same in all dimensions (except for the offer to receive mentoring). To examine how equivalent the groups were, we compared the groups' basic demographics and outcomes at baseline before the participant/treatment group youth received mentors. The equivalence tests for the participant versus comparison groups and the treatment versus control groups hold constant youth age, race/ethnicity, mode of survey administration (phone or in-person) and program. Table B.1 on the next page shows these comparisons. We find that the treatment and control groups are similar, differing at a 10-percent level of significance on only 3 of the 25 baseline variables. As is often the case, the participant and comparison groups are less similar, differing on 9 of the 25 variables.

When conducting a joint test on the similarity across all 25 characteristics, we again found more robust evidence of a difference between the participant and control/comparison groups, $F(30,777) = 2.65$, $p < .001$, than between the treatment and control/comparison groups, $F(30,515) = 1.40$, $p = .08$. Because tests of group differences at baseline were conducted on youth for whom we had follow-up data, the observed differences may be a result of differential attrition (that is, differences in characteristics of the youth from each group who dropped out of the sample due to lack of follow-up data), not only differences evident between the full groups of youth at baseline prior to attrition. However, as is noted below, we failed to find evidence of differential attrition.

Any variable that differed between the groups in either of these comparisons was included as a covariate in all outcome analyses to help ensure that we were comparing groups

Table B.1
Equivalence in Baseline Means Between the Treatment and Participant Groups and the Control/Comparison Group for the Analysis Sample

Scale/Indicator	Random Assignment Treatment Group (*n* = 308)	Quasi-Experimental Participant Group (*n* = 615)	Control/ Comparison Group (*n* = 321)	Difference Between Treatment and Control/Comparison Groups?	Difference Between Participant and Control/Comparison Groups?
Background/demographics					
Age	11.39	11.11	11.33	No	Yes
Male	58%	49%	58%	No	Yes
Ethnicity				No	Yes
White	42%	46%	39%		
Black	24%	22%	31%		
Hispanic	24%	19%	19%		
Other ethnicity	10%	13%	11%		
Risk profile				No	No
Hi Env-Hi Ind	29%	25%	27%		
Hi Env-Lo Ind	27%	22%	25%		
Lo Env-Hi Ind	11%	14%	14%		
Lo Env-Lo Ind	33%	39%	35%		
Mode of baseline survey administration (% by phone)	0%	26%	0%	—[a]	—
Mode of follow-up survey administration (% by phone)	100%	71%	100%	—	—
Outcomes					
Self-worth	3.16	3.24	3.16	No	No
Hope	4.29	4.41	4.23	No	Yes
Depression	1.56	1.45	1.56	No	Yes
Emotional symptoms[b]	1.74	1.70	1.71	No	No
Honesty	3.39	3.37	3.32	Yes	No
Prosocial behavior	3.56	3.58	3.55	No	No
Prosocial behavior[b]	2.52	2.60	2.54	No	Yes
Community service	.56	.55	.53	No	No
Social acceptance	2.64	2.70	2.62	No	No
Parent trust	3.37	3.43	3.38	No	No
Number of non-parental special adults	1.48	1.23	1.37	No	No
Non-parental help with college/jobs	.58	.57	.56	No	No
Substance use[c]	.04	.01	.04	No	Yes
Misconduct	.12	.11	.11	No	No
Legal problems[c]	.01	.01	.02	Yes	Yes
Conduct problems[b]	1.52	1.49	1.49	No	No
Self-perceptions of academic abilities	2.88	2.89	2.85	No	No
School liking	2.99	2.98	3.00	No	No
Educational expectations	3.29	3.38	3.28	No	No
Hours worked on homework	1.66	1.65	1.46	Yes	Yes
Skipping school[c]	.07	.04	.05	No	No
Grades	3.55	3.61	3.53	No	No

Note: Differences are tested at a .10 level of significance. When testing if outcomes differed at baseline across groups, we controlled for differences in the background/demographic variables listed in this table as well as the program to which the youth applied. The group means are all model-predicted means.
[a] Group differences on the measure were not tested because of a lack of variation in one of the groups.
[b] The outcome variable is reported by the parent; others are reported by youth.
[c] We tested effects in this area by assessing initiation of the behavior. Only youth who had not already engaged in the behavior were included in the analysis, and we tested whether, by follow-up, they had started engaging in the behavior.

that were as similar as possible—except for whether they had received mentoring. Specifically, in addition to age, race/ethnicity, program and survey mode, we controlled for the youth's: baseline scores on 7 (out of 22) measures of youth outcomes that showed evidence of differing ($p <$ .10) between either the treatment or participant group and the comparison group at baseline (that is, youth-reported depression, hope, honesty, substance use, legal problems and hours spent on homework, as well as parent-reported prosocial behavior) and baseline reports of 11 (out of 31) different indicators of risk that also showed evidence of participant/comparison group differences at baseline (that is, living in public housing, lack of a stable living situation, an unemployed parent, a sibling in foster care in the past five years, the youth having no close friends, gangs or illegal drugs in youth's neighborhood, single-parent status, youth belonging to a gang, frequent absences from school, youth failing or at risk of failing two or more classes and learning English as a second language). In addition, to obtain more precise estimates of the program's effects, we included the baseline value of the outcome measure as a covariate as well as indicators for individual and environmental risk status.[12]

Missing Data

Data were missing for small numbers of youth on the baseline measures of outcomes that were controlled for in analyses of program effects. To keep the sample as complete as possible, researchers have developed many different ways to handle covariate measures for which a given respondent did not complete the question (that is, "missing values"). We used the relatively straightforward technique of mean replacement. For those few observations that were missing on an outcome measure at baseline, we substituted the mean value of that variable for all the youth whose value was not missing. Because of how regression coefficients are calculated, this procedure ensures that the estimated coefficient on that variable is no different from the estimate that would have been calculated had the individual's data been omitted entirely.

Missing data for outcome measures at follow-up were not imputed in our analyses of program effects (that is, youth with missing data were excluded from the analysis for that outcome). The main reason for this type of missing data is that youth could not be found 13 months after their baseline to complete the follow-up survey. The overall response rate for the follow-up survey was 81 percent, which is quite good. However, technically, our results can be generalized only to the types of youth who responded to the follow-up survey.[13]

How do the youth included in the analysis differ from those for whom we lack follow-up data (that is, the "attriters")? Analysis of the baseline characteristics listed in Table B.1 revealed that those youth for whom we lack follow-up data were older (11.52 versus 11.24 years old), more likely to be Hispanic (44 versus 20 percent), less likely to be white (27 versus 38 percent), reported fewer depressive symptoms (1.47 versus 1.51), felt higher levels of social acceptance (2.76 versus 2.65) and were more likely to have reported misconduct (21 versus 13 percent) and having ever skipped school (22 versus 9 percent) at baseline. They were the same on all other dimensions. We also tested for differential attrition (that is, different patterns of attriter versus non-attriter differences for youth in the treatment and participant groups compared with those in the control/comparison group), but found no evidence of this when conducting joint tests across all baseline characteristics ($p = .26$ for treatment versus control/comparison group and $p = .29$ for participant versus control/comparison group).

Importantly, the attrition *rates* for the three groups of youth—the participant group, the treatment group and the control/comparison group—did not differ significantly [χ^2 (1) 3.09, $p = .21$]. The attrition rates were 19 percent for the treatment group, 17 percent for the control group and 21 percent for the participant group.

Multiple Hypothesis Testing

All statistical analyses run the risk of yielding a false positive result (incorrectly concluding that the program works when it really does not, known as a "Type 1 error"). The more tests one conducts, the greater the probability of finding, purely by chance, a statistically significant impact estimate when in reality there is no true impact. For example, if we test 10 independent and normally distributed impact estimates, we are likely to find at least one that is statistically significant at the $p < .10$ level simply by chance. One strategy for adjusting significance levels for multiple hypothesis testing is the Bonferroni adjustment, which establishes a statistical significance criterion by dividing the standard used in the study (for example, $p < .10$) by the total number of tests being performed. This strategy, however, is highly criticized for being overly stringent and severely increasing the likelihood that one would miss finding a true impact, declaring that the program is not effective when in fact it is effective (a "Type II error").

An alternative strategy that balances the risks of Type I and Type II errors better is the Benjamini-Hochberg (1995) family-wise adjustment (or "BH adjustment"), advocated by the What Works Clearinghouse.[14] This adjustment compares the p-values with an adjusted p-value criterion, q*(i/m) where q = the desired false discovery rate
   (in this report set to .10);
  m = the total number of p-values estimated; and
  i = the ith smallest p-value sorted from smallest
   to largest.

The statistical significance of a coefficient is then determined by organizing the p-values in ascending order for

all outcomes within a given domain (such as academic attitudes), comparing the largest p-value with the corresponding value of q*(i/m) and continuing up through the list until reaching the first p-value to satisfy the constraint that the estimated p-value < q*(i/m). All smaller estimated p-values are judged to be significant as well.

In Chapter 5, we examine how mentoring programs affect youth. What specific outcomes does it improve? Are there any specific negative consequences of participating in a mentoring program? Because, for these questions, we are interested in specific areas of influence, our tests are domain specific. Thus, the adjustment should be done separately for each family or domain of outcomes. However, as explained at the beginning of this appendix, we intentionally selected only one outcome in each of the 10 domains we examined (that is, emotional/psychological well-being, prosocial behavior and activities, peer relationships, adult relationships, antisocial behaviors, academic attitudes, academic behaviors, academic performance, outcome deterioration and outcome betterment), thus we do not need to perform any adjustments to these tests.

However, in addition to testing the exact nature of mentoring effects (for example, it improves this but may harm that), people may wish to test the more general hypothesis that "mentoring programs improve the lives of youth." Here, we are less interested in *what* outcomes are impacted than in whether the program has an impact on at least one of them. In this case we *should* perform this adjustment because in order to answer this question, we look to see if there are positive impacts on any of the 10 primary outcomes. In this case, negative effects are simply classified

as non-positive ones. Thus, for the purpose of examining this hypothesis, the domain-specific test used is a one-tailed test. (In the main text, we were interested in discovering both positive and negative effects; thus we used two-tailed tests.) Table B.2 indicates the BH-adjusted significance for the primary outcomes when we consider them all together as a test of the "mentoring has a positive effect on youth" hypothesis. (Only the outcomes that were significant without adjustment are listed.)

After applying this adjustment for multiple hypothesis testing we see that all of the quasi-experimental estimates, as well as all of the random assignment estimates, that were significant when unadjusted, remain significant. Thus, we can confidently conclude that participating in a mentoring program improves the lives of youth like those in our sample.

Analyses of Program Effects on Secondary Outcomes

While our primary analyses were restricted to just our 10 principal outcomes, we collected data on many more outcomes in hopes of further elucidating how mentoring might be affecting youth. Table B.3 on the next page shows the results of these analyses.

Subgroup Findings for the Primary Outcomes

As noted, to explore whether mentoring was more or less effective for particular types of youth, we examined whether estimated program effects varied significantly as a function of several background characteristics and experiences of

Table B.2

Statistical Significance Levels of Estimated Program Effects on the Primary Outcomes Using the Benjamini-Hochberg Adjustment (Using a One-Tailed Test)

Outcome	Effect Size (Standardized Mean Difference)	Estimated *P*-Value for a One-Tailed Test	BH Criterion i*[.10/10]	Adjusted Significance
Random assignment comparisons				
Number of positive changes	.15	.016	.01	Yes
Depression	.14	.018	.02	Yes
Quasi-experimental comparisons				
Depression	.32	.000	.01	Yes
Number of positive changes	.32	.000	.02	Yes
Social acceptance	.22	.002	.03	Yes
Grades	.19	.008	.04	Yes
Self-perceptions of academic abilities	.16	.019	.05	Yes

Table B.3
Estimated Program Effects on Secondary and Primary Outcomes

Scale/Indicator	Random Assignment Portion of the Evaluation	Quasi-Experimental Portion of the Evaluation
Emotional/psychological well-being		
Self-worth	.09	.20**
Hope	.02	-.04
Depression	-.14*	-.32***
Emotional symptoms[a]	-.19**	—
Prosocial behavior and activities		
Honesty	-.07	.07
Prosocial behavior	.00	-.02
Prosocial behavior[a]	.07	—
Community service	.01	.00
Peer relationships		
Social acceptance	.05	.22**
Adult relationships		
Parent trust	.08	.09
Number of non-parental special adults	.18*	.63***
Non-parental help with college/jobs	.06	-.31***
Antisocial behaviors		
Substance use[b,c]	.18	.42†
Misconduct[c]	.07	.04
Legal problems[b,c]	-.13	.20
Conduct problems[a]	-.13*	—
Academic attitudes		
Self-perceptions of academic abilities	.06	.16*
School liking	-.09	-.17*
Educational expectations	.07	.04
Academic behaviors		
Hours worked on homework	-.02	.13†
Skipping school[b,c]	-.21	.20
Academic performance		
Grades	.07	.19*
Number of outcomes showing positive change	.15*	.32***
Number of outcomes showing negative change	-.06	-.12

Note: The effect size for binary variables is calculated as ln(Odds Ratio)/1.65 per the WWC (p. 11 of http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_version1_standards.pdf). For count outcomes it is calculated as the difference in model-estimated means for treatments (or participants), and for the control/comparison group members it is divided by the observed standard deviation of the measure.

*** $p < .001$ level of significance.

** $p < .01$ level of significance.

* $p < .05$ level of significance.

† $p < .10$ level of significance.

[a] The outcome variable is reported by the parent; others are reported by youth. We did not conduct analyses with this measure in the quasi-experimental sample because we did not collect parent follow-up surveys in the quasi-experimental portion of the study.

[b] The analysis is limited to those youth who did not report this behavior at baseline.

[c] This variable is dichotomous.

youth. These characteristics and experiences, all of which were assessed at baseline, were as follows (all were assessed by parent reports except where noted):

- Gender;

- Ethnicity (white, African American, Hispanic);[15]

- Age;*

- Single-parent status;

- Living in poverty;

- Having a parent or other close family member who was incarcerated or had frequent problems with the law;

- Lacking a close friend;

- Experiencing potentially stressful life changes over the past year;*

- Quality of youth's relationships with parents and peers (based on youth report);[16]* and

- Presence of at least one special adult in the youth's life (based on youth report).[17]

Each of the characteristics noted with an asterisk (*) was represented as a continuous variable.

Examining possible differences in estimated program effects in relation to these youth characteristics across the 10 primary outcomes results in more than 100 comparisons. Statistics theory tells us that 10 percent of these comparisons could appear to be significant by chance alone. Thus, any significant findings should be considered only suggestive and only then if there is a theoretically plausible pattern of results.

In Table B.4, we report estimated program effects for youth subgroups when effect estimates varied significantly for a given outcome in relation to the relevant youth background characteristic or experience at a 10-percent or lower level of significance. (The significance level of this differential is noted next to the name of the outcome measure.) For example, the estimated program effect on parent trust for youth who did or did not have a parent or other close family member who was incarcerated (or had frequent problems with the law) differed at a .01 level of significance, with a significant estimated effect of .10 for youth without such a family member and a non-significant estimated effect of -.11 for youth with such a family member.

As noted in the report text, there is little evidence that effects differ by any of these background characteristics or youth experiences. However, we found two consistent patterns worth noting. First, estimated program effects on depression, parent trust and the number of outcomes showing negative change were all favorable and larger for youth who did not have either a parent or other close family

Table B.4

Significant Differences in Estimated Program Effects as a Function of Youth Background Characteristics and Experiences

Outcome	Program Effect Estimates (Unstandardized Regression Coefficients or Odds Ratios) for Different Youth Subgroups
Depression[†]	Incarcerated family member: Without -.11*** With .01
Parent trust**	Incarcerated family member: Without .10* With -.11
Number of outcomes showing negative change*	Incarcerated family member: Without -.25* With .28
Self-perceptions of academic abilities*	Age of youth:[a] Older .12* Average .06 Younger -.01
Prosocial behavior[†]	Age of youth:[a] Older .04 Average .00 Younger -.04
Skipping school* (odds ratio)	Has a close friend: Yes .60 No 2.70*
Grades*	Has a close friend: Yes .06 No -.18** Stressful life experiences:[b] Low -.01 Average .09[†] High .19**
Misconduct*** (odds ratio)	Living in poverty: Yes .65* No 1.82 Stressful life experiences:[b] Low 1.57* Average 1.12 High .80
Number of outcomes showing positive change	Stressful life experiences:[b] Low .17 Average .41** High .66***
Social acceptance	No significant subgroup differences observed

Note: The level of significance is noted for the test that effects were equal across subgroups. The effect estimates were derived from regressions that included the control variables noted in the text.

*** $p < .001$ level of significance.

** $p < .01$ level of significance.

* $p < .05$ level of significance.

[†] $p < .10$ level of significance.

[a] Younger (-1 standard deviation), average (sample mean) and older (+1 standard deviation) levels of youth age (in years) were 9.65, 11.24 and 12.83, respectively.

[b] Low (-1 standard deviation), average (sample mean) and high (+1 standard deviation) levels of stressful life experiences were .29, 1.68 and 3.07, respectively.

member who was incarcerated or had frequent problems with the law. Second, estimated program effects on grades, misconduct and the number of outcomes showing positive change were most favorable for youth who had experienced relatively high numbers of potentially stressful life experiences. These patterns and other differences found, however, could well be spurious. Additional research in this area is needed before firm conclusions can be drawn.

Appendix C: Who Were the Mentors?

Mentors bring different qualities to a relationship based on their own demographics (for example, gender, age) and past experiences. These characteristics may play an important role in how they approach the relationship and its chances for success. The volunteers' age (that is, whether they are adults or teens), for example, is key in determining how much youth benefit from mentoring in school-based programs. Herrera et al. (2008) found that adult mentors yielded much stronger benefits than their high-school-aged counterparts, particularly when staff did not provide the high school volunteers with additional supports (that is, high levels of communication).

In this appendix, we describe how the volunteers in this study were recruited and matched with youth. We also describe information collected in our follow-up surveys about the volunteers themselves—their previous experience with different groups of youth, and some of the struggles they faced in their past—to help contextualize our descriptions of the relationships these volunteers developed with their mentees and the benefits they were able to foster. We also use the risk profiles described in Chapter 2 to explore whether some of these characteristics may have shaped which youth these mentors were matched with.

Volunteer Characteristics

Almost half the volunteers who participated in the study were men. This makes sense given that, in community-based mentoring, same-gender matching is the norm,[18] and the study involved fairly equal proportions of male and female mentees. The roughly equal proportions of men and women who were eventually matched with youth, however, belie a number of challenges. As is typical in mentoring programs, services were requested more often for boys than girls, but more women than men volunteered, creating disproportionate numbers of boys on program waiting lists and challenges for programs in matching the women who volunteered. In fact, in one participating program, female volunteers experienced a long wait before they could be matched because the program enrolled so few girls.

Volunteers had an average age of 32 and were predominantly white (82 percent). The large proportion of white mentors reflects the broader volunteer population in these seven programs as well as volunteer demographics at a

---

Volunteer Recruitment

Participating programs reported using a range of strategies to recruit volunteers, including partnerships with businesses, colleges and faith-based organizations; participation in community activities, events or information sessions; and online strategies (for example, Facebook, Twitter). But word of mouth (for example, volunteers recruiting friends) was the strategy that, by far, agencies reported yielded the most mentors across the seven programs.

The programs also targeted recruitment efforts toward specific types of volunteers to help find mentors who "matched" the youth they anticipated serving. Five of the programs targeted male volunteers, for example through ROTC or university departments that tended to have a lot of male students (for example, engineering), implementing campaigns to seek out males and asking their male volunteers to encourage friends to apply. Five targeted mentors by race or ethnicity, and two specifically targeted gay and lesbian volunteers.

Over the course of the study, recruitment strategies did not change drastically, but five programs reported some changes. For instance, three began recruiting fewer younger (for example, college-age) mentors, noting challenges with maturity levels, inconsistency in meetings with mentees and difficulties with the time commitment required by the program.

---

national level. Many mentoring programs around the country struggle to recruit enough ethnic minority volunteers to serve the high numbers of minority youth who want a mentor who shares their ethnic background. This makes cross-ethnic matching the norm for minority youth, as was true in this study. Although 46 percent of volunteers shared the race/ethnicity of their mentees, most of these matches (89 percent) were white youth matched with white mentors. Only 9 percent of minority youth were matched with a mentor who shared their race/ethnicity.

The fairly high proportion of mentors who were college students (about a quarter of the sample) reflects the fact that in the university-based program, all volunteers were college students. Only a little more than half of all volunteers were currently employed.

Table C.1
Mentor Backgrounds

Mentor Characteristics and Experiences	Mentors in the Study
Background	
Average age	32
Male	48%
Employed	52%
College or graduate student	23%
Race/ethnicity	
African American	4%
Asian, Pacific Islander	1%
Hispanic	5%
Native American or Alaska Native	7%
White	82%
Other race/ethnicity	2%
Same race/ethnicity as youth[a]	46%
Previous experience with youth in different contexts	
Have children	15%
Have past mentoring experience	26%
Professional work with youth (for example, teacher, youth worker)	40%
Worked with youth in a volunteer setting	52%
Interacting informally with youth in their neighborhood (such as with relatives, babysitting or at church)	78%
No experience interacting with youth in any of these contexts	5%
Experience with different groups of youth	
Youth with behavioral, social or emotional difficulties	44%
Youth struggling academically	47%
Youth with family challenges	45%
Youth from diverse cultural backgrounds	45%
Youth living in poverty	35%
Youth involved with social services	27%
No experience with any of these groups of youth	32%

[a] Matches in which the "other" category for race/ethnicity was selected for either the mentor or youth were excluded from the calculation of this percentage because it was not possible in these cases to determine if the mentor and youth shared the same race/ethnicity.

We also asked mentors about their previous experiences with youth (before being matched in the program). Only 15 percent had children, but many brought other types of experience with youth to the program. About a quarter (26 percent) had previously been matched with another child in a mentoring program, and many had worked with youth in other formal volunteer and/or professional settings. Almost four fifths (78 percent) had informal experience interacting with youth (for example, relatives, babysitting, church). Only 5 percent had no experience interacting with youth in any of these contexts.

Many brought experience working with youth with behavioral, social or emotional difficulties (44 percent), youth from diverse cultural backgrounds (45 percent) and youth living in poverty (35 percent). (See Table C.1.) However, it is noteworthy that almost a third (32 percent) reported that they had no experience with *any* of the "at-risk" groups of youth we asked about—that is, the types of youth the programs were asked to recruit for the initiative.

Table C.2
Mentor-Reported Personal Challenges

	Percent Reporting Having Experienced This Challenge in Their Past
Individual challenges	
Health or mental health issues	13%
Academic problems	10%
Behavioral issues	9%
Personal drug or alcohol problems	6%
Personal legal problems	2%
Experienced at least one individual challenge	34%
Environmental challenges	
Family relationships	35%
Family drug or alcohol problems	22%
Lack of parental involvement	20%
Peer relationships	15%
Poverty	12%
Housing insecurity/homelessness	4%
Family legal problems	4%
Language or immigration issues	3%
Experienced at least one environmental challenge	49%
Experienced at least one individual and one environmental challenge	25%
Had not experienced any of these challenges	40%

In addition to asking mentors about their background and demographics, we also asked them whether they had ever personally experienced any of the challenges faced by many of the youth in the programs (see Table C.2). Very few (less than 15 percent in each of the six categories we asked about) had experienced "individual" challenges, such as academic challenges, health or mental health concerns, or behavioral issues; more reported experiencing environmental challenges, particularly in their relationships with parents, with at least 20 percent reporting a lack of parental involvement in their lives, family relationship problems and/or family drug or alcohol problems. Yet, only about 12 percent reported experiencing poverty (which the vast majority of participating youth experienced), and two fifths reported that they had not experienced any of the personal challenges we asked about. Thus, a large minority of the volunteers in this study had very little direct, personal experience with the kinds of issues the mentees in our sample were facing.

Making the Match

To understand how programs approached the process of matching these volunteers with participating youth, we asked staff in our program survey about their matching strategies. Responses suggest that staff emphasized geography and interests more than the youth's behavioral or social needs when making a match. That is, staff prioritized ensuring that mentors and youth could meet regularly and would have things in common more than trying to ensure that mentors could address the specific needs the child presented. (See Matching text box.)

To understand how mentors viewed themselves vis-à-vis the youth they were matched with, we asked them about the extent to which they felt they shared their mentee's personal background. Reflecting the relatively low proportions of mentors who reported facing the types of challenges faced by many of the youth, most mentors reported that they were "not very" or "not at all" similar to their mentee in the five areas we asked about (see Table C.3 on the next page).

Matching

During the initiative, participating programs used strategies they had used previously to match mentors with youth. In some cases, programs aimed to find similar characteristics between youth and mentors (for example, gender, language, ethnicity, interests), and in other cases, they tried to find characteristics or skills in the mentor that might complement the needs or interests of the youth (for example, matching a mentor who has experience with youth with behavioral problems with a youth who has such problems).

At the time the study began, the most common strategies were matching on gender (a matching criteria that, like language, is required, not simply considered, for most community-based matches) and hobbies/interests (on average, used for "almost all of the youth" within the programs); geography (used for more than half of the youth within the programs); age, race/ethnicity, youth's behavior problems or special needs (considered when matching about half of youth); and language, academic needs and sexual orientation (used for some, but fewer than half of matches). It should be noted that while most youth spoke English, when they did not, language was a primary consideration.

Over the course of the study, the programs did not drastically change their matching procedures. One expanded the mentee and mentor applications to provide staff with a better understanding of their interests. Another noted that staff began spending more time on matching and pulled in more seasoned staff to help create stronger matches. The university-based program began delaying matching until after the initial training session to provide staff with a better sense of the personalities of the mentors and the youth for whom they would be best suited. Understanding the potential importance of shared geography in determining match longevity, another program reported an increase in considering projections of locations rather than only where the mentor was living at the time of enrollment (for example, where she would be after college, if she was considering a new job, etc.). Another started having enrollment and match support staff work together before matches were proposed to try to foresee challenges that might occur in a given pairing.

We also wanted to explore more methodically the extent to which youth's needs seemed to play a role in determining which volunteers they were ultimately matched with. We did not assess whether programs consciously considered mentors' backgrounds during the matching process. However, we did examine whether mentor characteristics (for example, previous experience with youth) differed for youth with different risk profiles, based on the groupings we described in Chapter 2. In general, we did not find strong patterns to suggest that mentor backgrounds varied based on the risk level of the youth with whom they were matched. One exception was that volunteers with previous mentoring experience were more likely to be matched with youth who were high on only individual risk.[19] Perhaps staff felt that more seasoned mentors were better prepared for youth exhibiting these kinds of risks. However, previous experiences with specific groups of youth (for example, experience with youth struggling academically or those with behavioral difficulties) did not yield clear patterns in determining the type of youth with whom mentors were matched.[20]

Table C.3

Mentor-Reported Similarity with Youth in Background and Challenges Faced

	Percent of Mentors Reporting That Their Background Was "Not Very" or "Not at All" Similar to Their Mentee's	Percent of Mentors Reporting That Their Background Was "Somewhat" Similar to Their Mentee's	Percent of Mentors Reporting That Their Background Was "Very" Similar to Their Mentee's
Socioeconomic background	64%	27%	9%
Cultural background	52%	32%	17%
Family background (for example, number of parents or children in household, family dynamics)	67%	23%	9%
Social challenges	49%	39%	12%
Academic challenges	59%	30%	12%

Appendix D: Measuring Risk

To assess youth risk, we asked parents a wide range of questions at baseline about their child's background, challenges they might be facing at home or at school and the child's social and environmental surroundings. Their responses to these 31 questions contributed to our descriptions of youth risk and the risk analyses we conducted throughout the report.

Using Risk to Describe Youth

These questions fell into two broad categories: those that characterized environmental risk and those that characterized individual risk. "Environmental risk" considered challenges within the youth's surrounding environment that could predispose him or her to negative outcomes, whereas "individual risk" characterized challenges in the youth's behavior, social or academic functioning, or health. Both included three subcategories. Environmental risk included economic adversity, family risk/stress and peer difficulties; and individual risk included academic challenges, problem behavior and mental health concerns (see Items in the Risk Assessment text box).

In Chapter 2, Table 2.2, when we describe the youth involved in the study, we characterize youth as experiencing one of these six types of risk if their parent indicated that they had experienced one or more of the indicators within each of the subcategories. For example, we indicated that a child had experienced mental health concerns if the parent responded "yes" to either or both of the two questions listed under "mental health concerns" in the text box. And we characterized youth as experiencing either environmental or individual risk when parents noted at least one indicator within one of these two broader overarching categories.

The risk screening tool. All but seven of these questions contributed to a screening tool that we developed to help agencies quickly measure environmental, individual and "overall" risk as part of the enrollment process. Parents were asked to provide a "yes" or "no" response to 25 questions about their child and his or her surrounding environments. All of these questions are listed in the risk text box (those that are not starred) except for one item that was not retained because it was not accurately described by any of the risk categories we developed (that is, "One or more of the child's closest friends gets into serious trouble [for example, expelled, arrested, involved with illegal drugs].").

The first 14 questions in the risk screening tool asked about individual risk factors, and the last 11 asked about environmental risk factors (in our original risk assessment, the peer difficulties subcategory was included within individual risk). To be designated as "higher risk" with this initial screen, parents needed to indicate that the child experienced at least one individual risk factor and at least one environmental risk factor, and the total number of risk factors across the 25 items had to reach or exceed four. We later added the remaining seven items (starred in the text box) to our risk categories to enhance the study's assessment of environmental risk.

Creating the Risk Profiles and Using Risk as a Moderator

Throughout the report we also use risk as a "moderator" to test whether youth's level of risk was associated with other aspects of their experiences or outcomes in the program. For these analyses we needed to determine what we would designate as "higher" versus "lower" risk.

To create these groups, we first added the number of risk factors reported by parents in each of the six overarching areas (that is, economic adversity, family risk/stress, peer difficulties, academic challenges, problem behavior and mental health concerns) to yield six separate scores. We then divided those scores by the total number of risk factors possible in each respective area to yield the proportion of risk factors identified in each of the six areas. For example, a child whose parent reported that she had experienced four of the six challenges within the economic adversity group would receive a score of .67 for economic adversity. We then averaged these proportions across the three areas of economic adversity, family risk/stress and peer difficulties to calculate an overall "environmental risk" score. This approach, although not without limitations, is sensitive to both the number of risk factors present and the extent to which these were distributed across different domains of the youth's life. Thus, those scoring highest on environmental risk not only had a relatively large number of risk factors, but also had risk that was spread across several areas of the youth's environment. Similarly, we averaged the proportions of risk factors reported in the remaining three areas of academic challenges, problem behavior and mental health concerns to calculate an overall "individual risk" score. Finally, we examined the distributions of these scores and designated roughly the top half as being relatively "high" in environmental or individual risk and the bottom half as being relatively "low" in risk.[21] This process yielded the four risk profiles discussed throughout the report, that is: youth who were relatively

high on both environmental and individual risk; youth who were relatively low on both environmental and individual risk; youth who were relatively high on individual risk, but low on environmental risk; and youth were relatively high on environmental risk, but low on individual risk.

In analyses, we then used the resulting designations to test for moderation of program effects according to environmental risk status (high/low) and individual risk status (high/low), considering these categories both separately and in conjunction (that is, environmental risk X individual risk X mentoring group versus comparison group interactions).

Items in the Risk Assessment

All of the items listed below (except those seven that are starred) contributed to our initial risk assessment completed by parents at baseline. Parents responded either "yes" or "no" to each of these 24 statements (plus one additional statement that was ultimately dropped). The starred items were also included in the parent baseline but were not yes/no questions and did not contribute to our initial risk assessment.

Environmental Risk

Economic adversity

- The child lives in a public housing development.**
- The child lives in a situation where the parent believes the family could be forced to leave or evicted.**
- In the last 12 months, my family has experienced times when we had difficulty paying our bills.
- There are gangs or illegal drugs in the neighborhood where we live.
- Neither parent living with the child is currently working at a full-time job.**
- The child's family has a combined income below $20,000 (or receives food stamps).**

Family stress

- The child him/herself lives with a foster parent.**
- In the last five years, one or more of my child's siblings (or my child him/herself) has spent time away from home because he/she was placed in foster care.
- One or more members of my child's family struggles with alcohol or drug use.
- A significant member of my child's family (sibling, parent or other close relative) is in jail or prison or is often in trouble with police.
- My child lives with only one parent, guardian or other adult who takes care of him/her.
- My child has moved or changed where he or she lives two or more times in the last 12 months.
- The child's parents separated or broke up in the last year (for example, they started living in different places).**
- My child has seen or experienced many fights or arguments in our home in the last 12 months.
- My child has lost or lost contact with an important adult role model in the last 12 months (for example, a parent or other important adult died or moved out of our home).
- My child has experienced homelessness in the last five years.
- The child's parent reported that he or she did not complete high school.**

Peer difficulties

- My child doesn't have any close friends at school or in our neighborhood.
- My child has been picked on or bullied at school or in our neighborhood in the last 12 months.

Individual Risk

Academic challenges

- My child is currently failing or at risk of failing two or more classes/subjects in school.
- My child has a physical, emotional or mental condition that interferes with or limits his/her ability to do schoolwork at grade level (for example, ADHD, ADD or a learning disability).
- My child missed school often this past school year (three or more times a month).
- My child is learning English as a second language.

Problem behavior

- My child has used or experimented with drugs or alcohol.
- My child has been suspended more than once from school in the last 12 months.
- My child has been sent to juvenile hall or had contact with the police in the last 12 months.
- My child has run away from home in the last 12 months.
- My child belongs to a gang or spends time with gang members.
- My child often picks fights with other youth or bullies them.

Mental health concerns

- My child often says he/she feels alone, sad, upset, cries a lot or is unhappy.
- My child has been diagnosed with a mental health issue or is currently under the care of a mental health care provider (a therapist or counselor).

**Question was added to the initial risk screening tool.

Appendix E: Development and Validation of Mentoring Relationship Quality Scales[22]
By Daniel A. Sass and Michael J. Karcher

Surprisingly, there is only limited evidence of validity for most measures of mentoring relationship quality and inconsistent use of theory to guide the development of available measures (Nakkula, Harris in press). The purpose of the analyses described in this appendix was to further develop and validate youth-reported measures of the nature and quality of the mentoring relationship. We started with a pool of items from a few previously used scales as well as additional items that were developed as part of the current study. From this set of items, we sought to identify scales with good psychometric properties that could be used by researchers and practitioners to assess mentoring relationship quality.

We used as our organizing theory a framework for understanding mentoring interactions proposed by Karcher and Nakkula (2010). Their TEAM framework describes the focus of match activities and interactions (goal-directed or relationship-focused) as one of three core dimensions useful for understanding mentoring relationship styles. In this appendix, we explore the focus dimension of the framework, hypothesizing that our analyses would yield support for scales assessing goal-directed and relational foci as distinct but related aspects of match activities. Analyses were conducted using youth responses to 29 survey questions. These included 14 items used in previous P/PV studies: 5 of these items were used to form a scale measuring youth centeredness (Grossman, Johnson 1999); the remaining 9 items are those that Rhodes et al. (2005) proposed as measures of how much the mentor has helped the youth to cope (3 items) and the youth's feelings of unhappiness in the relationship (6 items). The other 15 items consisted of: 6 items developed by DuBois (2008) to assess a growth orientation or goal focus of mentoring interactions; 8 items developed for the study with the aim of measuring how much youth feel valued by their mentors; and 1 item measuring closeness (Herrera et al. 2007). All of these items were viewed as possible candidates for measuring either the relational or goal-directed focus of the TEAM framework.

Overall, analyses yielded support for using these items to form scales assessing youth centeredness and growth/goal-oriented engagement, respectively, as areas of focus for match interactions. Findings also provided support for the appropriateness of using the scales independent of the youth's gender, race/ethnicity or risk profile.

Method

Sample. The sample used in these analyses included youth in the study who were either in the treatment group of the randomized portion of the evaluation or the participant group of the quasi-experimental component of the evaluation, who were matched with a mentor and who completed a survey at follow-up ($n = 888$).

Data analyses. We first used exploratory factor analysis (EFA) to investigate the factor structure of the items (that is, to determine which items tended to reliably group together based on youth providing similar responses to them). These analyses were carried out separately based on two randomly selected halves of the overall sample ($n = 444$ each) in order to ensure that findings were not dependent on chance characteristics of the youth involved. Next, the two-factor model resulting from the final EFA with the first sample was assessed for how well that model worked with the data from the second random sample (n = 444) using a confirmatory factor analysis (CFA).[23,24]

We also were interested in whether the final confirmatory factor model resulted in a similarly good fit for youth from different groups (that is, males and females, youth belonging to different racial/ethnic groups and youth in the four risk profile groups that are the focus of this report).

Finally, we investigated the reliability of the scales that were derived from the factor analyses both for youth in each random half of the sample and for the same subgroups of youth for which model invariance was evaluated in the confirmatory factor analyses.

Findings

Exploratory Factor Analyses (EFAs). The exploratory factor analyses conducted with the full set of 29 items revealed support for a two-factor solution for each of the randomly selected halves of the sample.[25] The items loading most strongly on the two factors were consistent with the posited two-dimensional quality of the *focus* element of the TEAM framework, reflecting a goal-directed approach by mentors (named "growth/goal focus") as distinct from a relationship-building focus of mentors (for which we retained the name

"youth centeredness"). The former factor measures work toward an external goal, and the latter reflects the mentor's efforts to learn about the youth's interests and choose activities accordingly. Initially, some items loaded on both factors. Thus, items were removed one by one until an approximate simple structure (that is, items loading on primarily one factor, with small [λ< .30] cross-loadings) was obtained. This process substantially reduced the overall number of items from 29 to 10.[26]

The standardized factor loadings, along with the inter-factor correlations, are provided in Table E.1 on the next page. Based on item content, each item loaded (or correlated with) the appropriate content factor to which it was specified.[27] These analyses provide reasonable support for factorial/construct validity, with the exception of the rather high inter-factor correlation. Thus, ultimately, the original two measures of youth centeredness and growth/goal orientation emerged relatively intact and distinct, and were consistent with the TEAM framework characterization of interactions as either goal-directed or relationship-focused (Karcher, Nakkula 2010).

Confirmatory Factor Analyses (CFAs). The CFA conducted with the second random half of the sample to test the 10-item EFA-generated model described above provided a similarly good model fit, χ^2 (34) = 111.31, $p < .001$, CFI = .977, TLI = .970, RMSEA = .072, and practically ($\lambda \geq .30$) and statistically ($\alpha = .05/10 = .005$) significant factor loadings for all items, even after a Bonferroni adjustment (see Table E.1). Our invariance analyses further revealed that the two-factor, 10-item CFA model fit the data well for each gender and race/ethnicity of youth as well as for youth constituting each of the four risk profile subgroups.[28,29]

Scale Reliability Analyses. Results revealed satisfactory levels of internal consistency reliability (the correlation of items with other items in the same scale) for both the youth-centered and growth/goal focus scales within each of the random halves of the sample as well as for each subgroup of youth for which the invariance of model fit was investigated (see Table E.2 on page 116).

Summary and Discussion

The results of these analyses provide preliminary support for the psychometric adequacy of youth-reported scales of two facets of relationship quality with program-assigned mentors. It is encouraging that the item content of the two scales conform to aspects of mentoring relationship quality (youth centeredness and growth/goal orientation) that are well aligned with one of the core dimensions of a recent proposed framework for understanding mentoring relationship quality (Karcher, Nakkula 2010). The consistency of support for the scales across youth gender, race/ethnicity and risk profile group is also noteworthy as this suggests that it is appropriate for researchers and evaluators to utilize the scales with these differing groups, and that comparisons in scores on the scales across groups (for example, males and females) can be based on considerable evidence of validity.

The degree of overlap indicated between the two aspects of mentoring relationship quality is noteworthy for at least two reasons. First, mentors' incorporation of a goal/growth focus into their activities with youth has sometimes been viewed as inconsistent with actively considering youth interests and deciding on activities in a collaborative manner; our findings are more consistent with the view that these need not be viewed as competing approaches and, in fact, can often be characteristics of the same relationship (for further discussion, see Keller 2005). Second, because of the overlap that is evident between a goal/growth focus and youth-centered orientation, it will be important in future research to further establish the merits of distinguishing between these two aspects of relationship quality (for example, the extent to which scores on the two scales each contribute to prediction of outcomes such as match length after taking into account their association with each other).

More generally, these analyses are intended to be preliminary. Future investigations should explore the measurement of other dimensions of the TEAM framework as well as seek convergence across ratings of mentoring relationship quality from other perspectives, such as mentors and program staff.

Table E.1

Standardized Factor Loadings and Inter-Factor Correlations for the EFA and CFA Results Across the Two Random Samples

Item	EFA Sample 1		CFA Sample 1		CFA Sample 2	
	F1	**F2**	**F1**	**F2**	**F1**	**F2**
Youth centeredness						
1. My mentor almost always asks me what I want to do.	**.96**	**-.23**	**.75**		**.64**	
2. My mentor is always interested in what I want to do.	**.90**	-.01	**.91**		**.78**	
3. My mentor and I like to do a lot of the same things.	**.76**	.04	**.80**		**.79**	
4. My mentor thinks of fun and interesting things to do.	**.81**	.04	**.84**		**.95**	
5. My mentor and I do things I really want to do.	**.82**	.04	**.86**		**.83**	
Growth/goal focus						
6. My mentor and I spend time working on how I can improve as a person.	-.17	**.87**		**.70**		**.72**
7. My mentor helps me to set and reach goals.	.00	**.89**		**.87**		**.87**
8. My mentor and I work on projects together.	.18	**.58**		**.73**		**.73**
9. My mentor and I accomplish a lot of things together.	.22	**.65**		**.84**		**.91**
10. My mentor and I talk together about how to solve problems.	.16	**.68**		**.81**		**.78**
r (between-factor correlation)	**.73**		**.76**		**.78**	

Note: Statistically significant coefficients at $p < .05$ are bolded. However, the cross-loading onto factor two for item 1 was not statistically significant after the Bonferroni correction ($\alpha = .05/21 = .002$). F1 and F2 correspond to the youth-centeredness and growth/goal focus factors, respectively.

Table E.2
Internal Consistency Reliability Coefficients Across Groups

	Alpha	*M*-statistic	*P*-value
Youth centeredness			
Sample		.29	.59
Sample 1 (*n* = 444)	.82		
Sample 2 (*n* = 444)	.83		
Gender		4.38	.04
Male (*n* = 459)	.84		
Female (*n* = 429)	.80		
Risk factors		12.66	.005
Low on both types of risk (*n* = 330)	.80		
High on environmental risk only (*n* = 221)	.80		
High on individual risk only (*n* = 116)	.83		
High on both types of risk (*n* = 231)	.87		
Race/ethnicity		6.33	.10
White (*n* = 397)	.84		
Hispanic (*n* = 184)	.83		
African American (*n* = 197)	.82		
Other race/ethnicity (*n* = 110)	.75		
Growth/goal focus			
Sample		1.23	.27
Sample 1 (*n* = 444)	.84		
Sample 2 (*n* = 444)	.82		
Gender		1.22	.27
Male (*n* = 459)	.84		
Female (*n* = 429)	.82		
Risk factors		5.78	.12
Low on both types of risk (*n* = 330)	.83		
High on environmental risk only (*n* = 221)	.82		
High on individual risk only (*n* = 116)	.79		
High on both types of risk (*n* = 231)	.86		
Race/ethnicity		.23	.97
White (*n* = 397)	.83		
Hispanic (*n* = 184)	.83		
African American (*n* = 197)	.84		
Other race/ethnicity (*n* = 110)	.83		

Note: The *M*-statistic compares the Alpha (reliability coefficient) across groups within the relevant youth demographic category (for example, gender) to each other (for example, boys versus girls). After a Bonferroni adjustment (α = .05/12 = .004) per factor, there were no differences between category groups in scale internal reliability. Moreover, practically speaking these differences were rather small.

Appendix F: Analyses of the Effects of Rematching and Total Time Mentored on Youth Outcomes[30]
By David L. DuBois, Daniel A. Sass and Michael J. Karcher

Longer-term mentoring relationships have been posited to foster greater program benefits (Rhodes 2005). In a rigorous test of this hypothesis using data from P/PV's random assignment impact study of the BBBSA community-based mentoring program, findings suggested that program benefits became stronger as relationships persisted for longer periods of time over an 18-month follow-up period (Grossman, Rhodes 2002). By contrast, youth in relationships that terminated in less than three months showed declines in some areas (for example, self-esteem) relative to youth in the control group, thus suggesting that such short-lived relationships could be harmful. At the same time, there are also examples in the literature of relatively short-term mentoring experiences (for example, less than six months) being associated with favorable youth outcomes (for one example, see Wyman et al. 2010). Likewise, a recent meta-analysis of evaluations of youth mentoring program effectiveness failed to find an association between the average reported length of mentoring relationships and the estimated size of program effects (DuBois et al. 2011).

These mixed findings suggest the need to examine how the effects of the amount of time youth receive mentoring—that is, "total time mentored" (or "match length" for youth with only one match)—may be conditional upon other factors. Clearly, for example, it seems crucial to distinguish between shorter-term mentoring relationships that occur as an intentional part of a program's design and those that reflect a falling short of initially set expectations.[31] Also significant, however, even when longer-term mentoring experiences are the goal, may be whether or not during the course of their program involvement youth are matched with more than one mentor. Matching a youth with a new mentor after the youth's first mentoring relationship has ended ("rematching"), in fact, appears to be a fairly common practice in mentoring programs. It is noteworthy, therefore, that in a recent analysis of data from the P/PV impact study of the BBBSA school-based mentoring program, Grossman et al. (2012) found evidence of a *negative* impact on one of two academic outcomes examined (unexcused absences) for youth who, after their first match closed, were rematched with a new mentor during the same school year. The researchers noted that a number of processes could account for this finding, including a need for youth to have sufficient time to come to terms with the ending of a mentoring relationship, as well as the potential for the presence of a new mentor to draw attention to the first loss. They also stressed, however, the preliminary nature of their findings and the need for further study.

With this prior research as background, analyses were undertaken to examine the potential roles of both total time mentored and rematching in shaping the effects of mentoring program participation on the outcomes of youth in the present evaluation. The methodology and findings from these analyses are summarized in this appendix. In general, among those youth who were not rematched with a new mentor, we find little evidence of increasing benefits of program involvement as a function of the amount of time they had been matched with their mentor by the 13-month follow-up assessment. Interestingly, however, among youth who had been rematched, estimated program effects for certain outcomes dissipated with greater length of time mentored. In these instances, rematched youth with the most total time mentored tended to have the poorest outcomes relative to youth from the comparison group.

Method

Sample. The analyses reported in this appendix were based on those youth who had follow-up data ($n = 1,156$, including 835 youth who received mentoring and 321 youth in the control/comparison group).[32]

Analyses: propensity score matching. It is difficult to tease out accurate estimates of the actual effects of total time mentored and rematching on outcomes because youth characteristics can influence both outcomes and total time mentored/rematching (Grossman et al. 2011). Thus, it is always possible that youth characteristics (such as those that might lead a match to end early or for a youth to be rematched), rather than actual match "dosage" or the practice of rematching itself, are responsible for observed differences in youth outcomes. The current analyses make use of a statistical technique known as propensity score matching (PSM) to reduce the risk for this type of bias. PSM analyses are designed to improve one's ability to draw accurate conclusions from nonexperimental data by creating groups that are statistically equivalent (or as close to this as possible) on all known (that is, measured) characteristics at baseline. In the present context, PSM was used to estimate the effects that variations in total time mentored and rematching status had on mentored youth by identifying for each mentored youth a control/comparison-group youth who was well matched (that is, highly similar) to that youth on background characteristics and outcomes at the study's baseline assessment. This process was carried out separately for each of seven subgroups of youth defined by total time mentored and rematching status (see Table F.1 on the next page). Each control/comparison-group youth was eligible to serve as a "match" for a mentored youth in more than one of the subgroups.[33]

Table F.1

Sample Sizes for Total-Time-Mentored and Rematching-Status Groupings for Mentored Youth

	< 6 Months	6 to 11 Months	11 to 14 Months	14 Months or More
No rematch	No rematch < 6 months ($n = 137$)	No rematch 6 to 11 months ($n = 136$)	No rematch 11 to 14 months ($n = 269$)	No rematch 14 or more months ($n = 55$)
Rematched	Rematched < 6 months ($n = 20$)	Rematched 6 to 11 months ($n = 59$)	Rematched 11 to 14 months ($n = 24$)	—[a]

[a] Only five youth who were rematched received 14-plus months of mentoring. As a result, these youth were excluded from analyses, and there was no grouping for this combination of total time mentored and rematching status.

Total time mentored at the time of the youth's follow-up assessment[34] (which, in the case of youth who were not rematched, is equivalent to match length) was divided into four different categories: less than 6 months, 6 to 11 months, 11 to 14 months and more than 14 months.[35] Further distinguishing youth on the basis of whether they had been rematched with a second mentor resulted in the seven subgroups of mentored youth (and the group sizes) shown in Table F.1.[36]

Analyses: associations between total time mentored, rematching status and youth outcomes. To examine associations between total time mentored, rematching status and youth outcomes, we focused on the primary outcome measures examined in the report (that is, youth-reported measures of depression, prosocial behavior, social acceptance, parent trust, misconduct, skipping school, self-perceptions of academic abilities and grades, as well as the numbers of outcomes for which individual youth showed improvement or deterioration). Analyses controlled for baseline scores on the outcomes as well as all youth demographic and background variables that were included in analyses of overall program effects (see Appendix B) and three other primary outcome measures that exhibited evidence of association with the total time mentored and/or rematching status variables (that is, parent trust, social acceptance and misconduct). Analyses also controlled for whether the youth (or for control/comparison youth, the youth's matched counterpart) was in an active match at the follow-up assessment. In these analyses, we tested for differences in each outcome at follow-up in relation to total time mentored (that is, the four categories defined above) and rematching status as well as the interaction between them.[37] The latter test allowed us to examine whether the association between total time mentored and outcomes differed depending on the youth's rematching status (and vice versa). The total-time-mentored category was treated as a nominal variable (that is, it was not assumed that there would be a linear association between the total-time-mentored category and youth outcomes).[38]

Findings

Analyses revealed significant interactions between total time mentored and rematching status in relation to estimated program effects for four outcomes: parent trust ($p < .05$), self-perceptions of academic abilities ($p < .05$), grades ($p < .10$) and the number of outcomes with negative change ($p < .05$). Estimated program effects associated with each of these findings are summarized in Table F.2 on the next page. In general, the expected association between total time mentored and stronger estimates of program benefits was not evident. Notably, in fact, significant favorable program effects were concentrated for the most part among those who received less than 6 months of mentoring. Of further note is that for each of the four outcomes, there was an estimated program effect that reached or approached significance in the direction of harm, rather than benefit (that is, mentored youth scored less favorably than non-mentored youth from the control/comparison group). These effects were observed primarily among youth who had been mentored 11 or more months.

There also was a significant difference in estimated program effects in relation to rematching status for the number of outcomes with positive change ($p < .10$). Among youth who had not been rematched, a significant favorable program effect ($B = .69$, $p < .01$) was evident for this measure. In contrast, the estimated effect was weaker and non-significant ($B = .35$, $p > .10$) for those who had been rematched.

Summary and Discussion

Whereas conventional wisdom in the field and some previous research (for example, Grossman, Rhodes 2002) have suggested that longer mentoring relationships tend to facilitate stronger outcomes for youth, the present analyses failed to support this pattern. For 6 of the 10 outcomes examined, there was no evidence of an association between program effects and total time mentored. Moreover, for the remaining outcomes for which such differences were apparent, favorable program effects were concentrated among youth who

Table F.2

Estimated Program Effects (Unstandardized Coefficients) as a Function of Total-Time-Mentored and Rematching-Status Groupings for Mentored Youth

Outcome Measure	Rematching Status	Total Time Mentored			
		< 6 Months	6 to 11 Months	11 to 14 Months	14 or More Months
Parent trust	No	.03	.05	.09*	-.06
	Yes	.21[†]	.00	-.32*	—[a]
Self-perceptions of academic abilities	No	.12[†]	.07	.10*	-.01
	Yes	.23[†]	.11	-.17[b]	—
Grades	No	.18[†]	.16[†]	.14*	-.004
	Yes	.18	.17	-.30[†]	—
Number of outcomes with negative change	No	-.34*	-.11	-.08	.05
	Yes	-.75*	.26	.63*	—

Note: This table includes findings for only those measures for which we found a significant interaction between total time mentored and rematching status in predicting estimated program effects.

*** $p < .001$ level of significance.

** $p < .01$ level of significance.

* $p < .05$ level of significance.

† $p < .10$ level of significance.

a Only five youth who were rematched received 14 or more months of mentoring. As a result, these youth were excluded from analyses.

b $p < .11$

had relatively short-term mentoring experiences of less than 6 months. This aspect of our findings is somewhat perplexing. One contributing factor could be that in some instances mentoring relationships end relatively quickly as a result of the youth being on a favorable trajectory of development (for example, good grades in school, positive relationship with parent), such that there is a perceived lack of need or motivation to receive this type of support on the part of the youth. Under these circumstances, mentors too may be less inclined to invest their time and energy in continuing the relationship. Although speculative, as summarized in Chapter 3 of this report, it does appear that dynamics similar to these factored significantly in match closures.[39]

Of further note are the findings in which program effects in a harmful direction were apparent for youth who had experienced rematching with a new mentor. These results tended to emerge for rematched youth with the most total time mentored. Although preliminary, and in need of replication, this aspect of our findings cautions against programs using rematching as a strategy for achieving greater "doses" of mentoring for youth whose original matches have ended. Factors noted earlier that may complicate the rematching process, such as youth needing more time to process the ending of the previous relationship, may play a role in these findings. It could also be that some youth may not be good candidates for mentoring (at least as provided by the program or at the youth's given stage of development), such that simply adding more mentoring is of little value and even counterproductive. Likewise, should youth—even due simply to the "luck of the draw"—be faced with a second disappointing experience with a mentor, there may be heightened potential for internalizing that loss in ways that are harmful. The findings of the current analyses and limited prior research on this topic (Grossman et al. 2011) are in agreement in pointing toward the potential downsides of rematching youth with new mentors. Further investigation is clearly needed, however, to better understand the dynamics that shape the benefits and costs that may be associated with rematching under varying scenarios.

Appendix G:
Analyses of the Contribution of Case Managers to Mentor Support and Match Outcomes[40]
By Daniel A. Sass and Michael J. Karcher

The role that program staff, or case managers (CMs), play in the formation of mentoring relationships has received very little empirical attention in the mentoring literature to date even though the importance of mentoring program staff seems undisputed. To explore this assumption, we tested a theoretical model (see Figure G.1) in which CM competence, as rated by their supervisors, contributes to mentor perceptions of the quality of support received from the CM, and these perceptions, in turn, contribute to indicators of mentoring relationship quality as viewed by the youth—namely, the degree to which mentees feel close to their mentors (Parra et al. 2002) and report that their relationships include a goal or growth focus while also being youth centered—that is, that the mentor actively considers the youth's interests and takes a collaborative approach to selecting activities (Morrow, Styles 1995; Keller 2005). These latter indicators of match quality are posited to influence, in turn, the length of the mentoring relationship.

Overall, our results provide support for all of the hypothesized linkages in the model. However, as would be expected, based on our use of data from multiple reporters (that is, supervisors, mentors and youth) and the likelihood of numerous other influences on the variables in the model, the magnitude of associations linking model constructs was generally small; these linkages in most instances thus accounted for only a limited portion of the variation in measures that was observed across matches.

Method

Sample. Over the course of the evaluation, 1,156 youth received mentoring. Several of the mentors of these youth were supported by more than one CM. The analyses presented in this appendix focus on the role of the CM who supported each match in the study for the longest time. We focused on the length of youth's first match in these analyses to ensure that youth had the opportunity to be matched for the full length of the follow-up period (second matches did not have this opportunity). However, youth reports of relationship quality were for their most recent match (for example, if they had been matched twice, for their second match). Thus, to ensure that match length and reports of relationship quality were for the same match, 171 matches involving youth who were matched with more than one mentor during the study (that is, by the time of the follow-up assessment) were not included in the analyses. An additional 187 youth were excluded because their mentor did not have one CM for at least 80 percent of the time they were mentoring that youth. The final dataset consisted of

758 youth who had only one mentor and who had the same CM for at least 80 percent of the time they were matched. Seventy-three (73) CMs provided support to these matches. These CMs supported between 1 and 40 matches within their respective agencies (*median* = 7, *mean* = 10.38, *standard deviation* = 10.63).

Analyses. Because multiple mentors were supported by the same CMs (and thus the mentors' ratings of staff support were not all independent),[41] a multilevel structural equation model (SEM) was used to test our theoretical model.[42] This dependency was accounted for by nesting reports by youth, mentors and site supervisors within CMs. This captures their shared variance (as estimated using the intraclass correlation). All of these variables were utilized as level-one variables so that they could be considered unique predictors of match outcomes. Even supervisor ratings of CMs were considered at level one to allow them to serve as predictors of other variables (rather than including these ratings at level two, which would only explain to what extent the slopes and intercepts of the competence variable differed across CMs). Therefore, the principal level-one (or match-specific) variables include:

- The CM supervisor's judgment of the CM's overall competence in supporting the matches in general;

- Mentor-reported quality of support received from the CM;

- Two youth-reported measures of the quality of the interactions between the mentor and youth (youth centeredness and growth/goal focus);

- A one-item youth-reported measure of relationship closeness ("How close do you feel to your mentor?"); and

- Match length (at the time of follow-up, based on program data).

Mentors rated the support they received from program staff by responding to the following four statements (Karcher 2004) on a 1 to 5 scale ("Strongly disagree" to "Strongly agree"): "Program staff seem willing to help me," "Program staff have shared important information with me about my mentee," "Program staff have given suggestions on what I can do with my mentee" and "Program staff seem truly concerned about how well our match is going."

To assess CM competence, supervisors completed one set of questions for each CM, rating from 1 to 5 ("unsatisfactory" [1], "improvement needed" [2], "met expectations" [3],

"exceeded expectations" [4], "outstanding" [5]) the extent to which the CM (a) "effectively implemented feedback from her/his supervisor," (b) "actively sought feedback from her/his supervisor," (c) "positively and effectively engaged with team members" and (d) "worked to facilitate growth or improvement in areas where youth were in particular need of support," as well as (e) the extent to which the CM's "documentation (for example, case notes) effectively represented/captured each conversation with a match party (that is, mentor, youth, parent)." These prompts asked supervisors to think specifically about how the CM worked with matches involved in the initiative.

First, we conducted confirmatory factor analysis on the scales described above—that is, a model was fit with separate factors or latent variables for the variables shown in Figure G.1, which were: CM competence (X1); CM support quality (Y1); youth centeredness; growth/goal focus; and quality of mentor-youth interactions (Y2; a higher-order factor that reflected the contributions of both the youth centeredness latent variable and the growth/goal focus latent variable). Next, latent factor scores extracted from this model were used in the multilevel SEM model. In addition to the level-one variables listed above, appropriate steps were taken to control for variables that could bias our estimate of the true associations between latent factor scores.[43]

Findings

The overall fit of the hypothesized model, as assessed using multiple criteria, although generally good, also indicated a non-optimal fit and therefore that revisions to the model would be appropriate.[44] In particular, based on modification indices, a direct path from supervisor-rated CM competence to youth centeredness was added (see dashed path in Figure G.1). This model had acceptable fit according to all criteria (see endnote 44). The added pathway also makes theoretical sense in that regardless of how helpful and supportive mentors *viewed* their CMs to be, those CMs who were more competent in general, as reported by their supervisors, could better prepare mentors to be youth centered in their approach to their relationships with their mentees.

As can be seen in Figure G.1, all pathways in the final model were statistically significant (p < .05) and in expected directions, with the exception of a non-significant path from the quality of mentor-youth interactions to match length. The amounts of variance accounted for in each variable in the model by the corresponding predictor variables were as follows: mentor reports of CM support (R^2 = .013), closeness (R^2 = .011), quality of match interactions (R^2 = .006) and match length (R^2 = .220). These findings correspond to only small amounts of explained variance based on standards suggested by Cohen (1984; small ≈ .02, moderate ≈ .13, and large ≈ .26). Although prediction of match length from all three predictor variables produced a large effect size (R^2 = .220), all of the variables except closeness made rather small contributions to explaining match length. Collectively, supervisor ratings of CM competence explained only about 1 percent (R^2 = .013) of mentor reports of CM support, and mentor-reported CM support explained only about 1 percent of the variability in both closeness (R^2 = .011) and the quality of match interactions (R^2 = .006).

Summary and Discussion

Overall, these findings support the idea that more capable case managers are able to provide support that mentors experience as higher quality, and that both case manager competence and the support they provide likely contribute to higher levels of youth-centered and goal/growth-focused mentoring interactions, greater closeness experienced by youth and longer matches. Given that the measure of CM support quality was based on *mentor* reports, these results suggest that such benefits in match outcomes may arise at least in part because more capable CMs are able to support mentors in ways that mentors personally feel is helpful—specifically, that the mentors feel their case managers are concerned about the success of their matches, help the mentors learn more about their mentees and offer suggestions for match activities. In view of the lack of prior research directly assessing the competence of case managers in mentoring programs, and the relatively small magnitude of most of the associations that were found, the findings should be treated as preliminary and in need of further study. Ideally, future studies would obtain ratings of CM competence, mentor reports of CM support quality and youth reports of mentoring relationship quality using a longitudinal design in which these data are collected on a staggered timetable to better address the strength and direction of the causal pathways that are reflected in the model investigated in the current analyses. Still, as discussed in Chapter 6, the present findings point to promising directions for targeting training of staff in mentoring programs toward fostering behaviors that may help improve the quality and duration of matches.

Figure G.1
Test of Proposed Theoretical Model
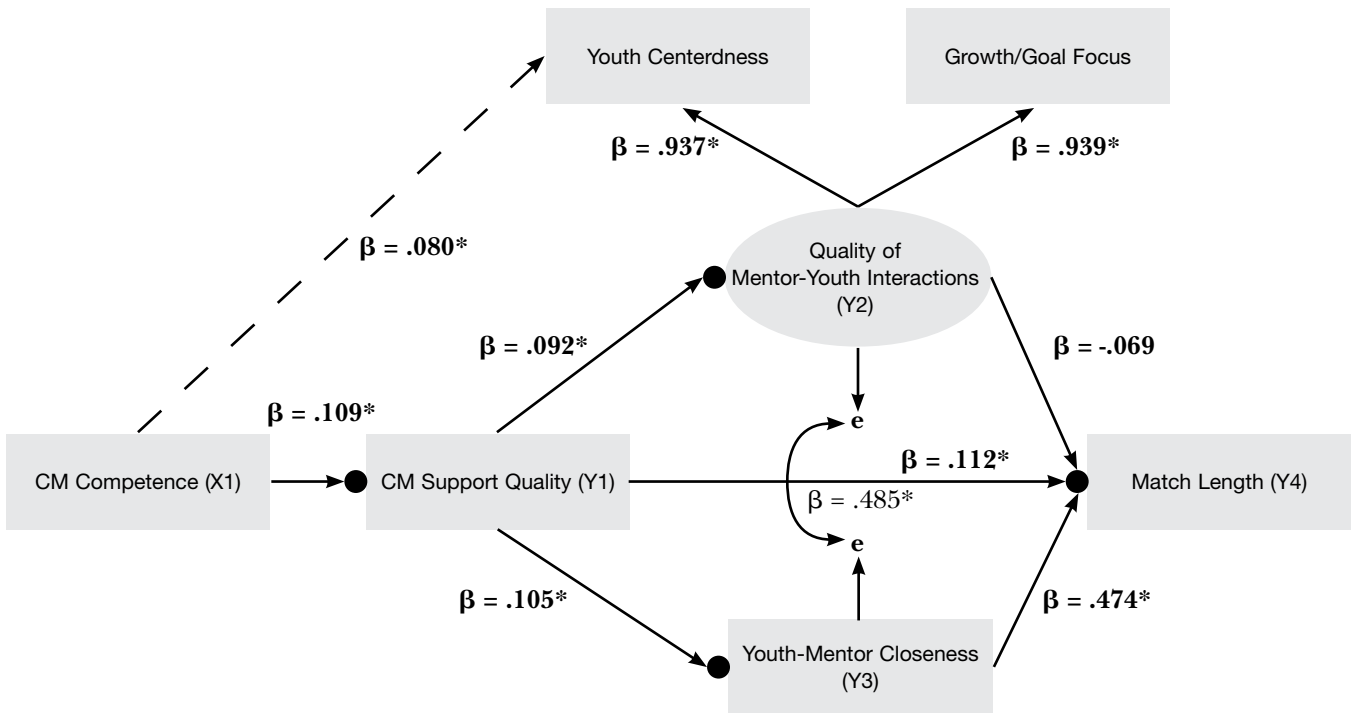
Note: This figure summarizes associations between variables of interest in our hypothesized model with the addition (dashed line) of one non-hypothesized model path.
   The β is a standardized estimate of the association between variables.
*Denotes a pathway that is statistically significant.

Appendices Endnotes

1. Agency records and feedback suggest that only a handful of families (less than 1 or 2 percent) who were asked to participate in either the impact or quasi-experimental portions of the study did not agree to participate.

2. At the time the study began, one of these two agencies offered activities to youth on its waitlist every couple of months (which would have also been offered to the youth in the control group). However, at that agency, less than 5 percent of the youth on the waitlist typically attended.

3. Of 321 youth in the control group for whom we were able to collect follow-up data, 7 reported having a special adult in their life who was a mentor from a program. This suggests that at least a small number of youth in the control group received formal mentoring from a source other than a program involved in the study.

4. Youth were required to have signed parent permission and youth assent forms to participate. Participation in the research study was not a condition for receiving a mentor. However, the parent consent form made it clear that study participants would have priority in matching, and if the parent chose not to participate, in all likelihood, his or her child would not be matched with a mentor until the study was completed.

5. These 231 youth are thus counted twice in this overall total (that is, 764 plus 777 equals 1,541 total). The total number of distinct youth who contributed to our baseline sample (without these "double-counted" control youth) is 1,310.

6. Agency records and feedback suggest that very few volunteers (less than 2 or 3 percent) who were asked to participate in either the impact or quasi-experimental portions of the study did not agree to participate.

7. This is a binary variable that is 1 if the youth reported participating in any of the specified activities in the past three months and 0 otherwise.

8. Analysis of this variable is limited to youth who did not report ever having skipped school at baseline.

9. We use Hageman and Arrindell's (1993) equation (equation number [4] in this appendix) to calculate reliable change indicators for each outcome. However, instead of using the raw observed post-test score, we used the predicted score from a regression including the full set of covariates and the baseline score on that outcome.

10. We excluded skipping school from this count because only youth who had not already skipped school were included in analyses using this measure—our analyses tested which youth began to skip school during the study period. Thus, improvement could not be assessed with this measure.

11. Approximately 1 percent ($n = 9$) of youth in the participant group in the quasi-experimental sample used for the analyses of program effects were never matched, and approximately 1 percent ($n = 3$) of the experimental treatment group used for these analyses were never matched.

12. Analyses for three of the outcomes (that is, substance use, legal problems and skipping school) were limited to youth who reported never having engaged in the behavior at baseline. Thus, for these outcomes, we did not include the baseline level of the outcome as a control.

13. In addition, for analyses of program effects on the youth-reported measures of antisocial behavior, we dropped those observations where youth "illogically" reported at baseline having engaged in a given behavior (for example, skipping school or having problems with the law), but then reported at follow-up that they had never exhibited these behaviors. These youth were excluded only from analyses examining the specific outcome measure for which their response was illogical. They were included in all other analyses and thus were not like the "attriters" who were excluded from all analyses because they lacked the entire follow-up survey.

14. What Works Clearing House. "What Works Clearinghouse Evidence Standards for Reviewing Studies (Version 1.0)." http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_version1_standards.pdf.

15. We included white, African American and Hispanic youth, omitting other ethnic groups from these analyses because the other groups were very small.

16. We created a combined measure of the quality of the youth's relationships with parents and peers by standardizing and averaging scores on the measures of parent trust and social acceptance and tested for both linear and curvilinear moderation of program effects by this measure, the latter being designed to be sensitive to the pattern of results found in earlier research (Schwartz et al. 2011).

17. Potential differences in program effects involving other subgroups, such as youth in foster care and those reporting legal problems (for example, arrest), were of interest but were not investigated due to less than 5 percent of the sample falling into such groups.

18. This was also true in this study; 96 percent of youth were matched with a volunteer who shared their gender. In all seven programs, at least 89 percent of matches were same-gender matches.

19. About a third (33 percent) of youth who were high on only individual risk were matched with mentors with previous mentoring experience. This proportion was significantly higher than those for youth in the lowest-risk group (23 percent) and youth high on only environmental risk (24 percent).

20. Responses to only two of the seven survey items for which we examined differences by risk group (that is, reports of experience with each of the six groups of youth we asked about as well as reports of having experience with none of these groups) showed significant associations with risk. First, mentors with experience with youth living in poverty were significantly more likely to be matched with the *lowest-* (37 percent

of this group) or highest- (37 percent) risk youth than to be matched with those youth who were high on only individual risk (28 percent of this group). Second, mentors with experience with youth involved with social services were more likely to be matched with the lowest-risk youth (29 percent of this group) than they were to be matched with youth high on only environmental risk (21 percent of this group). Perhaps mentors with this experience specifically asked not to be matched with youth experiencing more severe problems.

21. The percentages of youth designated as being high on environmental and individual risk were 48 percent and 40 percent, respectively. The remaining youth were designated as being at relatively low risk in these areas. Exact median splits (50-50) were not possible because of the nature of the distribution of each measure.

22. This appendix is intended to provide a brief overview of these analyses. A report with a more complete description is available from M. Karcher.

23. All statistical analyses (EFAs and CFAs) were conducted with Mplus 7 (Muthén, Muthén 1998–2012) using a weighted least squares mean and variance (WLSMV) adjusted estimator with a polychoric correlation matrix designed for ordered categorical data (4-point response scale).

24. The statistics employed to evaluate fit for both exploratory and confirmatory factor analyses were the robust χ^2, Comparative Fit Index (CFI), Tucker Lewis index (TLI) and Root Mean Square Error of Approximation (RMSEA). Based on Hu and Bentler (1999) and Marsh et al. (2004), TLI and CFI statistics greater than .90 and .95 reflect "acceptable" and "good" model fit and RMSEA values less than .08 and .06 are considered evidence of "acceptable" and "good" model fit, respectively.

25. Four eigenvalues exceeded 1 in each exploratory factor analysis. However, Horn's parallel analysis suggested a two-factor solution was equally appropriate and produced better model fit.

26. Several considerations—in particular, Horn's parallel analysis, the magnitude of eigenvalues, scree plot, model fit statistics and consistency with theory (that is, the hypothesized distinct youth-centered and growth/goal orientation aspects of the Focus dimension of the TEAM framework)—supported the two-factor solution with 10 items.

27. Given our concern about the high inter-factor correlations (suggesting a concern regarding discriminant validity) and the small cross-loading on factor 2 for item 1, a CFA was conducted with sample 1 also. The CFA results with sample 1 indicated that the cross-loading did not significantly impact the model fit statistics, χ^2 (34) = 98.59, $p < .001$, CFI = .983, TLI = .977, RMSEA = .065 or the magnitude of the inter-factor correlation (increased from .73 to .76).

28. Measurement invariance was indicated by non-significant $\Delta\chi^2$s for models testing invariance of the CFA fit across different subgroups of youth (for example, males and females).

29. Model fit indices for tests of invariance across gender, race and risk levels are available from M. Karcher.

30. This appendix is intended to provide only a brief overview of the analyses and findings. A more detailed report is available from D. DuBois.

31. Elledge et al. (2010), for example, reported preliminary evidence suggesting the potential effectiveness of a school-based mentoring program ("Lunch Buddies") in which youth, by design, receive a new mentor each semester.

32. As explained elsewhere in this report, a portion of the youth who were mentored in Year 2 of the study and for whom we were able to collect follow-up data ($n = 160$) had been control youth in the first year of the study. In general, to eliminate this source of non-independence in the data, when conducting propensity score matching for a given group of mentored youth, we excluded the control-year observations for those mentored youth who also had been members of the control group. Priority was given to retaining the Year 2 (mentored-year) observations for these youth to maximize the number of mentored youth within groups and thus statistical power. For one total time mentored/rematching status combination (6 to 11 months/no rematching), the number of mentored youth was relatively large and exceeded the number of youth in the control/comparison group. Accordingly, for the youth ($n = 70$) in this group who had both first- and second-year data, the first- (control-) year data were retained rather than the second- (mentored-) year data. A small number of observations also were excluded from analyses either because the youth involved was assigned to receive mentoring but was never matched ($n = 12$) or because the youth's mentoring experience (14 or more months of mentoring including a rematch with a new mentor) did not fit into any of the combinations of total time mentored and rematching status examined in the analyses ($n = 5$). For these reasons, the sample size reported in this appendix is less than the number of observations that were used in the main analyses of program effects for the study ($n = 1,244$).

33. Propensity score matching analyses were conducted using a program called MatchIt in R (Ho et al. 2007). In these analyses, youth demographics, background (for example, age, ethnicity, risk profile) and baseline measures of primary study outcomes were used as the basis for identifying matched pairs of mentored and control/comparison youth within each of the seven possible combinations of total time mentored/rematching status using nearest neighbor matching with a caliper of $.25\sigma_p$ recommended by Rosenbaum and Rubin (1985). The following steps were taken to assess the strength of the resulting matches: (1) visual comparison of the distribution shape of propensity scores for which suitable matches were identified with respect to whether there was a sizable overlap between the mentored and control/comparison groups; (2) tests of whether there was an approximately equal representation of both mentored and control/comparison group youth in each of the quintiles of the distributions; and (3) statistical tests for group equivalence on each variable used in the matching process. All indicated an excellent match between mentored and control/comparison youth within each of the seven total time mentored/rematching status groupings.

34. The follow-up was targeted to occur 13 months after the youth's enrollment in the study. However, there was frequently some delay in being able to obtain these assessments. Thus, it was possible for a youth to have been matched with his or her mentor for longer than 13 months by the time the assessment occurred. For this reason, the categories of total time mentored examined include durations of 14 months or longer.

35. Youth matched exactly 6, 11 or 14 months were placed in the higher of the relevant two possible categories.

36. Suitable matches in the control/comparison group were not found for 138 of the mentored youth. For this reason, the sum of the numbers of youth referenced in Table F.1 ($n = 697$) does not equal the total number of mentored youth in the study.

37. Each control/comparison group youth was assigned the same values of total time mentored and rematching status as the youth's matched counterpart in the mentored group. The specific tests of interest thus were whether group status interacted with total time mentored and/or rematching status.

38. As noted, control/comparison group youth could serve as matches for mentored youth in more than one of the categories shown in Table F.1. We accounted for this non-independence by modeling the correlations among observations for youth whose data were utilized in more than one of the categories, using maximum likelihood estimation and assuming an unstructured error process.

39. In keeping with these possibilities, among youth for whom mentor survey data are available, the percentage of youth with a mentor who reported, "My mentee didn't seem to need a mentor" as a reason for the relationship ending was higher among those with less than 6 months total time mentored (33 percent) than among those with either 6 to 11 months (16 percent) or 11 to 14 months (13 percent) total time mentored.

40. This appendix is intended to provide only a brief overview of these findings. A full description of the sample, measures, statistical analyses and results as well as a more extended discussion of the findings is available from M. Karcher.

41. The magnitude of the design effects of our level-one variables—which is directly related to the degree to which those mentors and youth who shared the same CM reported more similarly to each other and had more similar match lengths than those having different CMs—was consistently less than 2. This suggests that our tests of model pathways would not have been notably biased had we not conducted a multilevel structural model that accounted for this lack of independence (Muthén, Satorra 1995). However, to provide for a more conservative and technically appropriate approach to model estimation, we elected to fit a multilevel model that took into account the nesting of match-level variables within CMs. Ideally, we also would have accounted for potential non-independence of supervisor ratings of CM competence since in most instances a given supervisor provided ratings for more than one CM. However, the number of supervisors providing ratings was too small to support this additional level of model nesting.

42. The multilevel SEM analysis was conducted in Mplus 7 using maximum likelihood estimation.

43. The variables that we controlled for were: the mentoring program serving the match, mentor age and prior experience, youth age and gender, youth-reported depression, parent-reported conduct problems, and risk profile (as defined in the report). Following Newcomb and Bentler (1988), these variables were regressed on the latent factor scores, with the standardized residuals saved, and then converted to z-scores to ease the estimation and interpretation of the multilevel SEM model. It should be noted that step-wise regression analysis was used in this process; thus the latent factor scores were not necessarily residualized on the full set of potential control variables.

44. The statistics employed to evaluate model fit were the robust χ^2, Comparative Fit Index (CFI), Tucker Lewis index (TLI), Root Mean Square Error of Approximation (RMSEA) and Standardized Root Mean square Residuals (SRMR). Based on Hu and Bentler (1999) and Marsh et al. (2004), TLI and CFI statistics greater than .90 and .95 reflect "acceptable" and "good" model fit and RMSEA/SRMR values less than .08 and .06 are considered evidence of "acceptable" and "good" model fit, respectively. Using the standardized latent factor scores and observed scores adjusted for potential confounding variables and clustering by the 73 CMs, the model fit statistics (χ^2 (11) = 94.22, $p < .001$; CFI = .961, TLI = .892, RMSEA = .056, SRMR = .036) reflected inconsistent degrees of model fit. Specifically, the TLI, which accounts for model complexity, indicated a non-optimal fit, and therefore, based on modification indices, one additional path was added to the model. With the added path, the model fit was improved (χ^2 (10) = 47.06, $p < .001$), and all fit indices were acceptable or good (CFI = .982, TLI = .947, RMSEA = .048, SRMR = .026).