

University of New Orleans

ScholarWorks@UNO

University of New Orleans Theses and
Dissertations

Dissertations and Theses

Fall 12-20-2019

Evaluating the Effectiveness of a Community-Based Youth Non-Profit Organization at Increasing Prosocial Behavior and Decreasing Antisocial Behavior among Young Boys: A Pilot Study

Molly A. Miller

University of New Orleans, mamille2@uno.edu

Follow this and additional works at: <https://scholarworks.uno.edu/td>



Part of the [Child Psychology Commons](#), [Developmental Psychology Commons](#), and the [Policy Design, Analysis, and Evaluation Commons](#)

Recommended Citation

Miller, Molly A., "Evaluating the Effectiveness of a Community-Based Youth Non-Profit Organization at Increasing Prosocial Behavior and Decreasing Antisocial Behavior among Young Boys: A Pilot Study" (2019). *University of New Orleans Theses and Dissertations*. 2700.

<https://scholarworks.uno.edu/td/2700>

This Dissertation is protected by copyright and/or related rights. It has been brought to you by ScholarWorks@UNO with permission from the rights-holder(s). You are free to use this Dissertation in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you need to obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/or on the work itself.

This Dissertation has been accepted for inclusion in University of New Orleans Theses and Dissertations by an authorized administrator of ScholarWorks@UNO. For more information, please contact scholarworks@uno.edu.

Evaluating the Effectiveness of a Community-Based Youth Non-Profit Organization at
Increasing Prosocial Behavior and Decreasing Antisocial Behavior among Young Boys: A Pilot
Study

A Dissertation

Submitted to the Graduate Faculty of the
University of New Orleans
in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy
In
Applied Psychology
Applied Developmental Psychology

By

Molly A. Miller

B.S. University of New Orleans, 2012
M.S. University of New Orleans, 2016

December, 2019

Copyright 2019, Molly A. Miller

Table of Contents

List of Figures	iv
List of Tables	v
Abstract	vi
Introduction	1
Why are NPOs Important and What Do They Do?	4
How do Individual and Contextual Characteristics Influence Effectiveness?.....	13
What are the Barriers to Effective Evaluation and Why is Evaluation Important?	19
Aims and Hypotheses	26
Method	28
Participants.....	28
Procedure and Measures	29
Data Analysis Plan	43
Results	45
Aim 1	23
Aim 2	46
Aim 3	53
Supplementary Analyses.....	58
Discussion	60
Aim 1: Program Involvement	61
Aim 2: Program Involvement and Antisocial/Academic Outcomes.....	65
Aim 3: The Moderating Role of Individual and Environmental Factors.....	70
Limitations/Future Directions	74
Conclusion	80
References	81
Vita	92

List of Figures

Figure 1. Duration x Activity Participation Predicting Aggression-Mother's Report.....	51
--	----

List of Tables

Table 1. Reliability Estimates for Measures and Subscales of Study Variables	32
Table 2. Demographic Characteristics of Sample.....	39
Table 3. Group Means, Standard Deviations, and Range on Indicators of Main Study Constructs	40
Table 4. Bivariate Correlations of Main Study Constructs at Time 1	41
Table 5. Bivariate Correlations of Study Constructs at Time 2.....	42
Table 6. Hierarchical Regression Analyses Estimating Main and Interactive Effects of Program Involvement on Antisocial and Academic Outcomes at T1.....	50
Table 7. Hierarchical Regression Analyses Estimating Cross-Sectional and Longitudinal Effects of Program Involvement on Antisocial and Academic Outcomes at T2.....	51
Table 8. Hierarchical Regression Analyses Estimating Cross-Sectional and Longitudinal Main and Interactive Effects of Program Involvement and Duration on Antisocial and Academic Outcomes at T2.....	52
Table 9. Regression Analyses Estimating Main and Interactive Effects of Callous-Unemotional Traits and Program Involvement on Antisocial Outcomes	54
Table 10. Hierarchical Regression Models Estimating Interactive Effects of CU Traits and Program Involvement on Antisocial Outcomes at T2.....	55
Table 11. Regression Analyses Estimating Main and Interactive Effects of Exposure to Stressors/Trauma on Antisocial and Academic Outcomes at T1.....	57
Table 12. Hierarchical Regression Analyses Estimating Interaction Effects of Program Involvement and Trauma/Stressors on Antisocial and Academic Outcomes at T2.....	58

Abstract

Community-based youth non-profit organizations (NPOs) have become increasingly popular for the provision of youth prevention and intervention services, yet many youth NPOs lack the resources to undergo formal evaluation. Further, most existing program evaluations do not consider individual characteristics of the child or the child's exposure to stressors. The current pilot study sought to evaluate the extent to which boys participated in 1:1 mentoring and other program activities at the Son of a Saint (SOAS) NPO, an organization seeking to provide positive male role models for fatherless young boys. In addition, the current study examined the effects of program involvement on both prosocial (i.e., academic performance) and antisocial (i.e., aggression and delinquency) outcomes, as well as the moderating role of callous-unemotional (CU) traits and exposure to trauma/stressors on study outcomes. Data were collected from mothers ($N = 37$) and boys ($N = 27$) at the first assessment point, and from mothers ($N = 21$) one year later. Results of bivariate correlational and regression analyses at T1 indicated that boys who have been part of SOAS for shorter durations had higher levels of participation overall, and that behavioral/academic problems were associated with more program participation. Results at T2 indicated that participation in a greater variety of activities was related to lower levels of antisocial behavior. No significant interactions were detected for either CU traits or trauma exposure in the current sample. Implications of findings are discussed with regard to future program evaluation at SOAS. Detailed recommendations for overcoming the study limitations, particularly regarding the small sample size, are provided.

Key Words: Program Evaluation, Prevention, Juvenile Delinquency, Callous-Unemotional Traits, Aggression, Trauma

Introduction

For several decades, researchers have worked to identify aspects of the family and home environment that contribute to the psychological, behavioral, and social development of children. One particularly salient characteristic that often affects these aspects of development is the family structure (e.g., Gorman-Smith, Tolan, & Henry, 1999; see also McLanahan, Tach, & Schneider, 2013, for a review). For instance, children in single-parent families often experience more challenges than do children in 2-parent households (Gorman-Smith et al., 1999). Boys raised in homes without a father or stable male caregiver may be at increased risk for experiencing adjustment problems, including low academic achievement or engaging in aggressive behaviors or delinquency (Gorman-Smith et al., 1999; Hovee et al., 2009; Holmes et al., 2001; McLanahan, Tach, & Schneider, 2014; Simons, Wu, Conger, & Lorenz, 1994), for several reasons. First, the stressors associated with single parenthood increase the likelihood that mothers will use more negative parenting practices than two-parent households (Gorman-Smith et al., 1999). Negative parenting practices include harsh or inconsistent discipline, parental rejection, poor parental supervision, and parental neglect, and are associated with higher rates of aggressive and delinquent behavior among youth (e.g., see Hovee et al., 2009, Simons, Wu, Conger, & Lorenz, 1994, for reviews). Second, economic hardship as a result of relying on one parent's income to support a family is may require single parents to work multiple jobs to meet children's basic needs (e.g., food, shelter, clothing, etc.). Consequently, single parents often work longer hours, resulting in less time spent with families and less supervision of children, which in turn can increase risk of delinquency and poor academic performance (e.g., Holmes et al., 2001). Finally, young boys without a father or stable father-figure may be more susceptible to later antisocial behaviors than girls. In a recent review, McLanahan, Tach, and Schneider

(2014) reported that boys', but not girls', social-emotional development is negatively affected by father absence, specifically via increased externalizing behavior problems, and that the effect is strongest among boys whose fathers have been absent since early childhood. For many boys, the negative effects of father absence persist into adolescence, often in the form of increased risk-taking behavior such as associating with deviant peers and truancy (McLanahan et al., 2013).

Intervention programs often target fatherless boys in hopes of mitigating some of the risks associated with father absence. Son of a Saint (SOAS) is a mentorship-based community nonprofit organization (NPO) that seeks to improve the lives of fatherless young boys through emotional support, development of life skills, increased community involvement, and the formation of positive peer relationships. Son of a Saint seeks to alleviate some of this risk by pairing boys with mentors, who serve as adult male role models to help fill the need created by the loss of the father figure. Mentors also help coordinate boys' activities and involvement at various community events. Although the Son of a Saint program seeks to reduce risk for psychopathology among vulnerable youth in the community, SOAS has not evaluated their efforts to increase positive outcomes, such as academic performance, and prevent negative outcomes, such as antisocial behavior, among their participants.

In recent years, the number of community-based, nonprofit organizations (NPOs) providing social services has grown so much that NPOs now outnumber state- or publicly-funded organizations responsible for administering youth prevention and intervention programs (e.g., Bach-Mortensen & Montgomery, 2018; Norris-Tirrell, 2014). However, many NPOs do not undergo formal evaluation of their policies or activities and fail to monitor the effectiveness of their program goals (Bach-Mortensen & Montgomery, 2018; Chalmers, 2003). Evaluating the effectiveness of prevention programs is critical to identify which aspects of an intervention are

working and which are not, as well as for whom the intervention is or is not working. Without this information, program directors and staff are unable to refine the intervention efforts to ensure that, at a minimum, youth are not being inadvertently harmed by program practices.

The primary goal of the proposed study is to conduct a pilot evaluation on the effectiveness of the Son of a Saint NPO in New Orleans, LA. The current study used data collected from both mothers and boys currently participating in the SOAS mentoring program. The data included information on boys' aggression, deviance, and delinquency, level of neighborhood danger, exposure to stressors, and academic performance. Boys who are more engaged in SOAS programming were expected to engage in less deviance and aggression and to perform better academically than boys less engaged in the SOAS programming. Consequently, data was collected regarding the dosage of exposure to programming. Analyses were conducted for both mother- and child-report of antisocial behaviors and academic performance, as well as for mother-reported program involvement and academic/antisocial outcomes one year after the initial assessment.

The following sections will begin with a discussion of why and how NPOs seek to prevent antisocial behavior and increase prosocial behavior among youth. Next, the individual (i.e., personality traits) and contextual (i.e., exposure to stressors and/or trauma) characteristics of children and their environment that may influence the effectiveness of interventions or services provided by NPOs will be described. Finally, the importance of evaluating NPOs and the challenges faced by NPOs and evaluators that may act as barriers to evaluation of program effectiveness will be discussed.

Why are NPOs Important and What Do They Do?

The most effective programs for increasing positive adjustment, while reducing negative outcomes, are those that target multiple domains of youth functioning, including family relationships, school attendance/performance, peer relationships, and community involvement. For example, Multisystemic Therapy (MST; Henggeler & Bourduin, 1990; Henggeler, Schoenwald, Borduin, Rowland, & Cunningham, 1998, 2009) is considered a “model” program for reducing the incidence of delinquent behavior among youth with conduct problems. (e.g., aggression and delinquency; Greenwood & Turner, 2011; NJJ, 2011). The overarching goal of MST is to keep adolescents with a history of behavioral problems in school, at home, and out of trouble (Henggeler & Schaeffer, 2010). A team of two to four licensed clinicians provide therapy in the home and work with parents, children, and schools to set up and maintain a treatment plan for the adolescent; treatment plans often include up to 60 or more hours per week of direct contact with the therapist (Henggeler & Schaeffer, 2010; Henggeler et al., 2009). Therapists are available to make home visits at any time of the day or week to accommodate parent schedules and crisis situations. Treatment goals vary according to the needs of each family, but often include parent training, helping families to develop support systems, and addressing issues such as parental substance use and other sources of environmental stress (Henggeler et al., 2009). MST has been extensively tested in a wide range of settings and delivery sites. For example, Timmons-Mitchell and colleagues (2006) found that, at 18-month follow-up, youth who received MST had fewer arrests, better school performance and engagement, and demonstrated significant improvements in their behavior towards others, including respect for the rights of others and conformity to laws compared to the treatment as usual (TAU) group. Similar findings have been

reported in numerous other effectiveness studies on MST (e.g., see Markham, 2017 for a review).

However, despite the effectiveness of MST and similar programs for reducing negative and increasing positive outcomes, such programs are impractical for most families within the community. First, like most evidence-based treatments (EBTs), MST requires strict adherence to therapeutic protocols to maintain its effectiveness. Meta-analytic evaluations of MST have found significantly weaker treatment effects when community mental health workers comprised the team of therapists than when the team was comprised of more closely-supervised clinical psychology graduate students (e.g., Curtis, Ronan, & Borduin, 2004). Thus, MST requires a substantial time commitment from families who may see little to no improvement in their children's academic, social, behavioral, or family functioning if therapists are not closely supervised. Second, implementing complex, evidence-based interventions and ensuring strict adherence to treatment protocol is often expensive. For example, the estimated first year costs of MST are around \$933,000, including costs associated with hiring and training personnel, supplies, overhead, and licensing costs (Dopp, Borduin, Wagner, & Sawyer, 2014). Further, continued monitoring of treatment fidelity, training, and therapist salaries result in an average annual cost around \$7068 per family (Blueprints Programs, 2011). Funding of this magnitude is usually provided at the state- or federal-level and can be difficult and time-consuming to obtain, particularly for those communities most in need (Dopp et al., 2014). The absence of adequate resources to implement model programs is one of the primary factors contributing to the increased reliance on community-based NPOs for addressing the needs of vulnerable youth (Bach-Mortenson & Montgomery, 2018). Finally, MST and similar programs are often designed to treat children and young adults who have already displayed significant behavioral and

academic problems, including severe aggression and involvement with the criminal justice system (see Markham, 2017, for a review). However, there are many children and adolescents within communities who can benefit from services like those provided in MST before academic and behavioral problems arise. For these children, NPOs like Son of a Saint fill the critical need for community-based prevention programming that seeks to provide opportunities for academic/personal enrichment (e.g., receiving tutoring, participating in community service), promote prosocial behaviors (e.g., developing positive peer relationships, learning conflict resolution skills), and discourage antisocial behaviors such as aggression and delinquency.

Academic performance and school commitment. Children who do well in school are more likely than lower-performing students to develop supportive relationships with peers, teachers, and other adults (see Farrell et al., 2007). These positive relationships help children develop a sense of belonging or connectedness to the school environment, increasing their motivation for academic achievement (Resnick et al., 1997). In contrast, children who underperform in elementary, middle, or high school are less likely than better-performing students to complete high school or attend college, which may lead to difficulties becoming self-supporting young adults (see Felson & Staff, 2006). Further, poor academic performance can lead children to feel less connected to school, which has been associated with poor social adjustment (i.e., peer relationships) and school avoidance and truancy (e.g., Birch & Ladd, 1997).

Poor academic achievement also has been associated with increases in aggressive and delinquent behavior from elementary school through high school, although the relationship between academics and behavior problems is complex (e.g., Brier, 1995; Maguin & Loeber, 1996; Tremblay, Masse, Perron, & LeBlanc, 1992; Wasserman et al., 2003). For example,

children who exhibit disruptive behavior problems from an early age are likely to perform poorly on tests of academic achievement at all grade levels (Tremblay et al., 1992). For these children, disruptive behavior may interfere with their ability to attend to teacher instructions and to behave appropriately in the classroom such that poor behavioral regulation undermines academic performance. In contrast, other children may become discouraged by early academic failures, leading them to develop a negative self-image, which can interfere with the development of prosocial peer relationships (e.g., Felson & Staff, 2006; see also McEvoy & Welker, 2000). For these children, behavior problems may develop as a way of coping with peer rejection and the disappointment or disapproval they receive from family or teachers (e.g., Zamora, 2005).

Regardless of whether behavior problems precede or result from academic difficulties, substantial evidence suggests that intervention programs that improve academic performance also lead to reductions in behavior problems (Najaka, Gottfredson, & Wilson, 2001). Specifically, Najaka and colleagues (2001) conducted a meta-analytic review assessing the impact of academic-focused prevention activities on behavior problems (e.g., substance use, aggression, criminal activity, and defiance of authority) and found that positive changes in school performance, commitment, and attachment were consistently associated with reductions in problem behavior. This effect remained significant regardless of whether children's behavior problems began before or after entering school, indicating that any programming or interventions that lead to academic improvements are highly likely to reduce the incidence of behavior problems among youth.

Prevention of antisocial behavior problems. Although rates of juvenile arrests have declined over the past twenty years, youth under the age of 18 comprised nearly 10 percent of all arrests made in the United States in 2017 (OJJDP Statistical Briefing Book, 2017). The financial

burden associated with juvenile delinquency is considerable. From arrest to adjudication to incarceration, a single juvenile offender can cost taxpayers nearly \$47,000 per year in the state of Louisiana (Petteruti, Schindler, & Ziednberg, 2014). In contrast, community-based prevention programs can save taxpayers an average of \$7- \$10 for every dollar invested (Drake, Aos, & Miller, 2009). While not as easily quantifiable, juvenile delinquency also creates considerable psychological, social, and emotional costs to both offenders and victims. For example, juvenile offenders often do not complete high school, struggle to find gainful employment as adults, and are likely to have continued involvement with law enforcement once they reach adulthood (e.g., see Petteruti, Walsh, & Velazquez, 2009). Ongoing criminal activity increases the likelihood that the individual will spend considerable portions of his adult life incarcerated, which can negatively impact relationships with family, employment potential, and emotional well-being (e.g., ongoing stress associated with criminal justice system involvement or prison conditions). Costs to victims can include monetary, personal, or property losses, loss of life, and psychological consequences associated with the emotional stress of victimization (Drake et al., 2009). Thus, preventing the onset of delinquency and other serious behavior problems can reduce both the financial and psychological cost of juvenile offending within communities, making such programs appealing to policy makers, potential funding sources, and the community as a whole (Greenwood & Turner, 2011).

NPOs often use a mix of strategies to reduce or prevent behavior problems. For example, the Keep Safe program (Kim & Leve, 2011) is a community-based NPO that targets social skills-building and self-confidence among middle school-aged foster children and parents. Results of several randomized control trials (RCTs) indicated that Keep Safe is moderately effective for reducing behavior problems and substance use, although more studies are needed to determine

long-term efficacy of the strategies used (Kim & Leve, 2011; Kim, Pears, Leve, Chamberlain, & Smith, 2013; Smith, Leve, & Chamberlain, 2011). Girls on the Run (GOTR) is an NPO designed for preteen girls, which seeks to promote social competence and healthy lifestyles through physical activity (see Fredricks, Naftzger, Smith, & Riley, 2017, for a review). Whereas Keep Safe is specifically considered a delinquency prevention program, GOTR, like Son of a Saint, falls under the more general “positive youth development” umbrella. Evaluation studies on the effectiveness of GOTR have yielded mixed findings. The annual evaluations available on the GOTR website reveal strong positive effects on overall self-esteem and engagement in healthy behaviors (e.g., Gabriel, DiGiacchino DeBate, High, & Racine, 2011; Riley & Weiss, 2015). However, at least one study found no change in target behaviors (Ames, 2013). While both Keep Safe and GOTR seek to promote positive developmental outcomes for participants in similar age groups, they do so using very different strategies, targeting children at varying risk levels. Keep Safe and Girls on the Run are two examples of the hundreds of community-based NPOs currently operating around the country and illustrate the variable nature of youth organizations. Son of a Saint combines delinquency prevention efforts, like the Keep Safe program, with positive youth development efforts, like the Girls on the Run program. However, an additional and key strategy used by the Son of a Saint program is mentoring. Mentorship-based prevention programs have become increasingly popular since promising outcomes of the Big Brothers/Big Sisters (BBBS) were first reported over twenty years ago (Grossman & Garry, 1997).

Currently, mentoring is one of the more widely-used strategies among prevention programs with over 5000 programs currently using this approach in some form (see Tolan et al., 2013). Although mentoring in general is a broad and variable concept, there are some commonalities across mentoring programs. One common feature of mentoring is that children

are paired with an older, more experienced individual who provides guidance and support to youth on a variety of issues, including encouraging prosocial behaviors and social responsibility (Dubois et al., 2011; Karcher & Nakkula, 2010). Additionally, NPO mentors are most often volunteers, rather than employees, which suggests that they are intrinsically motivated, either by the mentor's own life experiences or a sense of duty to his or her community (Rhodes, 2002; see also Tolan et al., 2013). Potential drawbacks to the use of volunteer mentors is that volunteers may not be prepared or trained to handle more challenging youth (Grossman & Tierney, 1998). One way that SOAS attempts to counteract this limitation is by requiring mentors to be regularly involved with the program and to interact with the boys at various community events and activities for at least 3 months before being paired with specific mentees. This practice also helps to ensure that volunteer mentors are both committed and well-suited to the role they have agreed to play in boys' lives. Finally, as is the case with SOAS, nearly all mentoring occurs as part of a program with multiple components, including community service, tutoring, life skills training, or group activities (Aos et al., 2004). Therefore, it is important for evaluators to consider the effects of mentorship within the broader framework of the NPO's programming to determine whether and how mentoring contributes to the effectiveness of the overall program.

Despite these common features of mentorship-based programming, there is substantial variation in the frequency (i.e., dosage) of mentor-mentee time and goals of the mentor-mentee relationship across various programs. Some mentor-mentee relationships are based more on developing specific skills (e.g., academic, social, and vocational skills), while others are more generally focused on providing youth with a positive role model who can offer guidance and emotional support as needed (Hart-Johns, Courser, & Kirk, 2012). The amount of time mentors spend with mentees also varies considerably, including within and across NPOs. Most programs

request that mentors spend a minimum number of hours per month with mentees, but ultimately it is the decision of the mentor-mentees to determine how often meetings take place (see Tolan et al., 2013).

Importantly, mentoring-based strategies hold promise in preventing or reducing antisocial behavior and increasing prosocial behavior among youth, particularly when used in conjunction with other strategies (e.g., increasing community involvement and academic performance, life and vocational skills training; Lipsey, Howell, Kelly, Chapman, & Carver, 2010; Hart-Johns et al., 2012). In a recent meta-analysis, Tolan and colleagues (2013) found that youth at risk of behavior problems (i.e., youth at risk due to exposure to neighborhood crime, a history of disruptive behavior, or parental criminal activity) who participated in any mentoring program showed reductions in aggressive behavior and delinquency as well as increases in academic achievement. Effect sizes were stronger when mentoring practices included high levels of emotional support or advocacy for the youth in various settings (e.g., school, social services, etc.).

Similarly, Tierney and Grossman (1995) evaluated specific outcomes among youth (ages 10-14) participating in BBBS programs compared with youth not participating in BBBS. At the 18-month follow-up, youth involved with BBBS were 46 percent less likely to initiate substance use and 32 percent less likely to have engaged in physical violence. Furthermore, BBBS participants missed half as many days of school compared to the control group and showed statistically significant improvements in academic confidence, grades, and overall quality of family relationships. Although specific aspects of mentoring strategies that led to these improvements were not examined, mentors and mentees had high levels of contact during the

study (i.e., 70% met with mentors at least 3 times a month for more than 3 hours at a time), suggesting that the frequency of contact may have contributed to the positive outcomes.

Additionally, Keating, Tomishima, Foster, and Alessandri (2002) considered the impact of mentoring on reducing behavior problems among children (aged 10-17) with a history of behavioral or academic problems. Youth who spent at least 3 hours a week with mentors over a 6-month period showed statistically significant reductions in both internalizing (i.e., anxiety and depression) and externalizing (i.e., aggression, inattentiveness, delinquent activities) behaviors compared with the waitlist control group. Taken together, mentoring frequency likely influences the impact of the mentoring experience on youth adjustment such that spending more time with mentors contributes to the effectiveness of mentorship-based programs.

In contrast, other studies have found that mentorship-based programs failed to reduce behavior problems or, in a few instances, children's behavior problems worsened over the course of their involvement with the program (Royse, 1998; Slicker & Palmer, 1993). For example, Royse (1998) found no significant improvements from baseline assessments on measures of self-esteem, GPA, school absences, or disciplinary problems in a sample of youth after 13 months of mentoring. Several possible explanations were posited to explain the null findings, including poor record-keeping by mentors, which complicated the assessment of time spent with mentees, and the fact that the majority of mentees were from low income households and resided in high crime neighborhoods. Possibly, mentoring alone is not sufficient for improving the overall quality of life for children living in environments in which they are regularly exposed to crime and violence. Slicker and Palmer (1993) examined differences between effectively and ineffectively mentored youth, quantifying effectiveness by the amount of time mentors spent with mentees each week. Youth in the study were identified by teachers and parents as being at

high risk of school dropout and behavior problems. Among those youth in the ineffective mentoring group, only 69% returned to school the following year, compared to 100% of the effectively mentored group. Further, those in the ineffective mentoring group showed poorer academic performance compared to both the effectively mentored and control group and experienced significant declines on measures of self-esteem and self-concept over the course of the study. Notably, studies that reported null or negative findings were evaluating the effects of mentorship as a standalone intervention strategy. That is, no other intervention strategies (e.g., community service or skills training) were provided to youth receiving mentoring. This pattern supports findings from Tolan and colleagues' (2013) meta-analysis that mentoring is more effective when used in conjunction with other strategies.

Taken together, findings from evaluation studies of mentorship-based prevention programs suggest that mentoring can be among the more effective methods of preventing behavior problems among high-risk youth, provided the mentees receive frequent and effective mentoring for the duration of their involvement with the program. Findings that under certain conditions, some children may exhibit *more* problem behaviors following interventions highlight the importance of evaluating the effects of programs and organizations which aim to reduce or prevent youth behavior problems. However, certain individual and contextual characteristics of children and their environments

How do Individual and Contextual Characteristics Influence Effectiveness?

Individual characteristics. Aggressive behavior during childhood is perhaps the most studied predictor of later behavior problems, including delinquency (e.g., Dodge et al., 2006; Loeber, 1990; Valois et al., 2002). Although a comprehensive review of the aggression research is beyond the scope of the proposed study, certain types of aggressive behavior are differentially

related to intervention effectiveness and are therefore worth noting here. Specifically, aggression can be distinguished into two categories: reactive and proactive (see Crick & Dodge, 1996, Marsee et al., 2011; Poulin & Boivin, 2000 for reviews). The primary distinction between types is the motivation underlying the aggressive act. Reactive (sometimes referred to as “hostile”) aggression typically occurs in response to provocation or perceived threat or harm (e.g., Dodge, Lochman, Harnish, Bates, & Pettit, 1996). Reactively aggressive children often exhibit signs of emotion dysregulation, including low frustration tolerance and high impulsivity (see Card & Little, 2006). Reactively aggressive children tend to attribute hostile intent in ambiguous situations (Crick & Dodge, 1996; see also de Castro, Veerman, Koops, Bosch, & Monshouwer, 2002 for a review), often resulting in rejection by peers and social isolation (Poulin & Boivin, 2000). In contrast, proactive (or “instrumental”) aggression is goal-oriented, driven by a need for personal gain and often requires little or no provocation (e.g., Dodge, 1991; Vitaro, Brendgen, & Trembley, 2002). Proactively aggressive children view aggressive behavior as an effective means of achieving goals and are therefore less likely than reactively aggressive children to be distressed by aggressive acts (Dodge et al., 1997). Proactively aggressive children are less likely to be rejected or victimized by peers and are typically more skilled in social situations compared with reactively aggressive children (e.g., Dodge & Coie, 1987).

Complicating the assessment of aggression even further, the consequences of proactive aggression may vary based on temperamental characteristics of children. Children who demonstrate more callous-unemotional (CU) traits are more likely to have a severe and stable pattern of antisocial and delinquent behavior over time (e.g., Frick & Dickens, 2006; Frick & White, 2008; Hawes, Price, & Dadds, 2014). Considered a downward extension of adult psychopathic traits, CU traits are characterized by a shallow affect (i.e., showing or experiencing

little to no emotion), lack of guilt or remorse, manipulation, and disregard for the rights and feelings of others (e.g., Frick, Cornell, Barry, et al., 2003). Children with more CU traits typically exhibit more proactive aggression and seem to be less sensitive to punishment than children with fewer CU traits (Hawes & Dadds, 2005; Marsee & Frick, 2007). This relatively small subgroup of antisocial youth are responsible for most of the delinquency with very poor treatment outcomes (Frick, Cornell, Bodin, et al., 2003; Hawes et al., 2014). CU traits have only recently been recognized as a clinically significant form of conduct disorder, namely conduct disorder “with limited prosocial emotions,” in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association, 2013). Many community-based prevention programs are unaware of CU traits. As discussed previously, the likelihood of treatment resistance and the unique risk for stable and severe delinquency, CU traits may minimize intervention effectiveness and should therefore be examined within the context of community-based youth NPOs. To date, little is known about whether CU traits differentially impact outcomes of youth participating in community programs such as SOAS.

Environmental characteristics. Residing in dangerous or socially/economically disadvantaged environments can increase children’s risk for behavior problems in part because these environments have high levels of community violence and trauma. Exposure to community violence and other stressful or traumatic events during childhood has been linked to increases in violent behavior and nonviolent delinquency during adolescence (see Miller & Marsee, 2019, for a review). Several mechanisms have been explored that may contribute to this relationship. First, growing up in a neighborhood characterized by frequent or severe violence teaches some children that violence and aggression effectively manages conflict (e.g., Guerra & Huesmann, 2004). Consistent with Bandura’s (1973) social learning theory, these children quickly learn that

aggression and deviance are a means to gain respect and, by extension, survive in their immediate environments (Anderson, 1999). Relatedly, widespread community violence may simply become normalized for children who must witness violent events on a regular basis (Schwab-Stone et al., 1995). Finally, witnessing or experiencing traumatic events can interfere with the development of emotion regulation skills (e.g., Shields & Cicchetti, 1998; Tull, Barrett, McMillan, & Roemer, 2007). Trauma-exposed children who cannot regulate or control intense negative emotions are more prone to perceive ambiguous situations as threats and to engage in reactively aggressive behavior (e.g., Shields & Cicchetti, 1998; Eisenberg, Guthrie, et al., 2000; Eisenberg & Spinrad, 2004). NPOs like SOAS seek to mitigate this risk by pairing boys with emotionally supportive mentors, exposing boys to positive experiences and activities within the community and by teaching alternative strategies for resolving conflict. Thus, many aspects of SOAS and similar programs may benefit youth exposed to higher levels of neighborhood danger or other traumatic stressors, as these children are likely to be more in need of developing healthy coping skills and support systems compared to peers exposed to lower levels of violence or trauma.

Exposure to frequent stressors or trauma during childhood may also place youth at higher risk of academic difficulties compared to peers with less frequent exposure to trauma or stressors (e.g., Ammerman, Cassisi, Hersen, & Van Hasselt, 1986; Holt, Finkelhor, & Kantor, 2007; Larson, Chapman, Spetz, & Brindis, 2017; Overstreet & Mathews, 2011; Slade & Wissow, 2007). For example, Holt and colleagues (2007) found that multiple victimization experiences were related to lower GPAs among a sample of 689 youth aged 10-12. The authors posit that this association may be due in part to the substantial social difficulties reported among participants in the multiple victimization group, although no analyses were included to further explore the

contribution of social functioning to academic performance among participants. While the strength of the association between trauma and academic performance often varies depending on the type (e.g., peer victimization, physical/sexual abuse, community violence exposure, etc.) of stressor and the age or ages of youth at the time of the traumatic event, studies consistently demonstrate that exposure to multiple types of trauma or stressors is a strong predictor of academic difficulty (see Overstreet & Mathews, 2011, for a review). Furthermore, a large body of neurobiological research has shown that exposure to trauma and stressors during childhood can cause long-term changes in the development of neural systems that regulate emotional and cognitive responses (see Anda et al., 2006; Larson et al., 2017; Majer, Nater, Lin, Capuron, & Reeves, 2010, for reviews). For example, Slade and Wissow (2007) found that the relationship between childhood trauma and adolescent academic performance was partially attributable to cognitive deficits, particularly deficits related to attention and memory. Thus, exposure to early trauma or stressors can result in differences in the physiological development of neural systems that can potentially affect both cognitive and emotional functioning over time. Given that all boys at SOAS have experienced the trauma of losing a father to death or incarceration, it is likely that some boys participating at SOAS have experienced substantial trauma and/or stressors, which may negatively influence boys' ability to focus and remember material learned in school. As described below, there is ample evidence that programs and interventions designed to enhance resilience among vulnerable youth can reduce the influence of negative life experiences on developmental outcomes.

Protective factors as intervention targets. Prevention programs often seek to identify individual, familial, social, and environmental characteristics that reduce the likelihood that vulnerable children will engage in delinquent behaviors later in life (e.g., Coie et al., 1993; Kim

et al., 2015). These characteristics are often referred to as “protective factors,” and are generally defined as any behaviors or attributes that buffer individuals at risk from negative outcomes (e.g., Najaka et al., 2001; Rutter, 1987). Individual characteristics that seem to shield children from developing maladaptive behaviors include the ability to tolerate frustration, negative attitudes towards aggression and delinquency and younger age (Jolliffe, Farrington, Loeber, & Pardini, 2016; Loeber & Farrington, 2001; Pardini, Loeber, Farrington, & Stouthamer-Loeber, 2012). At SOAS, mentors can provide guidance to help boys handle frustration in more productive ways (i.e., not lashing out aggressively in response to perceived hostility or threat), and can model negative attitudes towards aggression, violence, and rule-breaking. The program also requires that boys join SOAS between the ages of 9-10, a developmental period during which social and emotional behavior is more malleable than it will be once boys reach adolescence (e.g., Cantor, Osher, Bergm Sleyer, & Rose, 2016). Protective factors related to children’s environment include having a positive school climate, having a non-deviant peer group, and a sense of connectedness to the community (Bernat, Oakes, Pettingell, & Resnick, 2012; Pardini et al., 2012). SOAS is highly active within the community and mentors strongly encourage boys to participate in community-based activities, such as attending professional sporting events, and engaging in community service. Boys also attend group mentorship sessions with other boys in the program, during which they are taught skills such as etiquette, teamwork, time management, and integrity. In addition to skill-building, the group mentorship sessions provide boys with an opportunity to develop prosocial relationships with peers. Further, boys have opportunities for scholarships that provide tuition assistance at local private or parochial schools. For many of the boys attending SOAS, receiving a scholarship to a private school substantially improve their academic opportunities since private schools often have lower

student-to-teacher ratios and more access to academic resources, such as textbooks, technology, and teachers, compared to local public or charter school systems.

Evaluation studies that consider individual and contextual aspects of a children's environment that may influence susceptibility to NPO interventions are clearly needed. Evaluation studies provide program staff with critical information regarding the needs of participants and areas of program improvement, such as providing staff with additional training in trauma-based interventions or placing more emphasis on trauma-specific protective factors. The final section will address challenges faced by NPOs and evaluators that can impede effective evaluation practices. The section will conclude with a discussion on the importance of overcoming these barriers so that ongoing program evaluation can take place.

What are the Barriers to Effective Evaluation and Why is it Evaluation Important?

Community-based NPOs often use a mix of strategies to promote change in functional domains, such as academics, behavior, or social competence. Numerous resources exist for the development and implementation of programs using evidence-based “best practices.” For example, the Office of Juvenile Justice and Delinquency Prevention (OJJDP) provides a Model Programs Guide that includes a list of prevention and intervention programs deemed “effective,” “promising,” or “no effects” and includes links to the effectiveness studies supporting these classifications (“OJJDP Model Programs Guide”, 2019). In addition, the OJJDP offers implementation guides for several evidence-based treatments (EBTs) and links to program developer websites, which often include additional resources, such as manuals and funding recommendations. However, most NPOs are not required to use these resources and, as many NPO service providers have no training in EBTs, many are not aware that such resources exist (Wandersman & Florin, 2003). As a result, programming is often unlikely to be based on

empirically evaluated evidence. The gap between scientifically-informed interventions and the practices implemented by NPOs can create barriers for both NPO providers and evaluators (see Bach-Mortensen & Montgomery, 2018).

Barriers for NPO providers. Two types of barriers have been identified that are likely to interfere with an NPO's ability to conduct evaluations. The first is inadequate funding (e.g., Bach-Mortensen & Montgomery, 2018; Murray & Tassie, 1994). Many well-established community-based NPOs (e.g., Big Brothers/Big Sisters, Boys & Girls Clubs of America) have financial resources to undergo extensive evaluation (Bach-Mortensen & Montgomery, 2018; Murray & Tassie, 1994; Roth & Brooks-Gunn, 2016). Large or nationally-based organizations also are more likely than smaller NPOs to receive federal or state funded grants that require regular monitoring and reporting (Foster et al., 2009). Small, locally based NPOs, like SOAS, rely most heavily on donations from local businesses and private citizens (Foster et al., 2009; Murray & Tassie, 1994). Funds are allocated based on need, such that most funding will be used to add program resources, increase salaries or improve recruitment efforts (Bach-Mortensen & Montgomery, 2018). Evaluation studies are not cheap and can cost upwards of several thousand dollars for small scale evaluations (Love et al., 2016). As a result, smaller NPOs will not have sufficient financial resources to engage in rigorous, independent evaluation

The second type of barrier for NPOs is a lack of resources (including time and staff) and skills to develop and implement evaluation procedures. Often, NPOs rely heavily on volunteer staff, most of whom are not trained mental health or evaluation professionals and who have other time commitments (i.e., full- or part-time jobs). Evaluations can be time-consuming, as they require ongoing monitoring by NPO staff of program practices and participant outcomes (e.g., Carman & Fredricks, 2008; Despard, 2016; Herman, 1998). As such, evaluations often result in

more administrative work, both for volunteers and paid program staff, and can be difficult for program directors to ensure that all volunteers and staff adhere to tracking and monitoring policies at all times (Bach-Mortensen & Montgomery, 2018; Love et al., 2016). The directors and staff at SOAS have indicated that they are highly committed to being active participants in developing an evaluation strategy. Most recently, the organization developed a smartphone app the boys must use to sign up for events and activities, which allows participation data to be easily collected and tracked with little to no additional time requirements for volunteers or staff.

Barriers for evaluators. Even when NPOs are able to overcome these barriers, evaluators still frequently encounter several challenges as well. First, many NPOs have somewhat broad mission statements or goals, such as to promote positive youth development or to provide youth with the necessary life skills to become functioning members of society once they reach adulthood. These goals can be difficult to operationalize and measure, making the evaluation process even more complex and challenging, as it is not always clear whether programs are meeting their intended goals (e.g., Murray & Tassie, 1994). Second, if program staff are unwilling to adapt programming based on evaluation results, the process can waste time and resources for both evaluators and NPOs (e.g., Love et al., 2016; Murray & Tassie, 1994). A third related challenge is that organizations may not have the ability to make changes to programming (Love et al., 2016). For NPOs utilizing the maximum resources available (e.g., funding, staff/volunteers, and facilities), changes can be difficult to make. In other words, translating research into practice can be difficult for NPOs without the financial means to implement changes. Thus, program evaluators must partner with NPOs to communicate their confidence in the results and to help make recommendations that are within the capacity of the NPO to implement. This challenge is likely especially salient for small NPOs who may not have

the ability to employ individuals with knowledge of evidence-based practices (Norris-Tirrell, 2014; Roth & Brooks-Gunn, 2016; Wandersman & Florin, 2003). A final challenge for evaluators is educating program staff on the implementation of best practices and the importance of regular evaluation with rigorous monitoring. NPO staff and interventionists are not required to receive formal training to operate NPOs and, as previously noted, programming is often based on the intuition and experiences of directors and staff members. While this does not necessarily result in poor interventions, program directors and NPO staff may be unaware of the potential for null or iatrogenic effects that certain strategies may have on youth. Without regular evaluation and assessment of youth behavioral, academic, and social outcomes, NPOs cannot refine strategies to maximize the benefits of the program for youth while minimizing any risk. Thus, NPOs must not only be provided the tools and information necessary to conduct or undergo evaluation procedures, but also demonstrate that they are willing to make changes to program strategies if and when recommended.

Importance of evaluation. Although community-based NPOs comprise a large portion of social services and delinquency prevention efforts, little is known about the effectiveness of these organizations and the services they provide. Evaluation is particularly important for organizations targeting youth at risk of engaging in antisocial behavior, given that there have been several documented instances of peer-based interventions that increase risk of engaging in antisocial behaviors (e.g., Dishion et al., 1999). That is, a lack of awareness of current academic research on recommended best practices may place NPOs at an even greater risk for iatrogenic practices than randomized control trials, thereby increasing rates of delinquency rather than ameliorating delinquency risk. Even well-developed, widely-used interventions and organizations have reported null or negative findings after consistent, rigorous, evaluation of

their efficacy or effectiveness (e.g., Petrosino et al., 2003; West & O’Neal, 2004). For example, the Scared Straight program, which gained popularity in the 1970s and 1980s, was designed to deter youth from engaging in delinquent behaviors by depicting the harsh reality of prison life by bringing youth on tours of prisons and allowing them to interact with inmates (Finckenauer, 1982). A similar program, the California SQUIRES (San Quentin Utilization of Inmate Resources, Experience, and Studies) added a component in which the youth are shown pictures depicting graphic prison violence (Lewis, 1983). Programs like Scared Straight and SQUIRES are appealing to communities because they are inexpensive and fit within the “common sense” notion that the threat of prison would be a sufficient deterrent to youth criminality (Finckenauer, 1982; see also Petrosino et al., 2003). However, evaluations of such programs demonstrated that the intervention strategy was ineffective for preventing or reducing future arrests (e.g., Finckenauer, Gavin, Hovland, & Storvoll, 1999; Lundman, 2001; Sherman et al., 1997). For example, after evaluating his Scared Straight program, Finckenaur (1982) reported that 41% of youth in the treatment group went on to commit new criminal offenses, compared to only 11% of the no-treatment control group. The SQUIRES program also found iatrogenic effects such that youth receiving the intervention had a higher re-arrest rate (81%) within the first 5 months post-treatment than the and control group (67%). Although Finckenaaur did not examine potential mechanisms by which the negative results may have occurred, Lewis’ (1993) evaluation of SQUIRES found that youth in the treatment group expressed significantly more positive attitudes about criminal behavior following the intervention, both compared to the no-treatment group and to their own baseline scores on the same measure. Scared Straight and SQUIRES represent examples of programs designed based on “common sense” rather than evidence and highlight the importance of evaluating both program outcomes and specific program strategies to determine

which aspects of the interventions may be contributing to the outcomes. Iatrogenic effects also have been found in interventions involving a high level of contact with more deviant peers in the program and low staff supervision (Arnold & Hughes, 1999; Cecile & Born, 2009; Dishion et al., 1999). Although many of the boys at SOAS do not have a documented history of severe behavior problems, the influence of one or two deviant or aggressive boys can be significant, particularly if those boys are viewed as more popular than the boys who embrace the prosocial goals of SOAS (Dishion et al., 1999).

While iatrogenic effects are cause for concern, iatrogenic effects are rather rare in intervention studies (e.g., Lochman, Dishion, Boxmeyer, Powell, & Qu, 2017). More often, interventions are found to have no effect on the targeted outcomes. Project D.A.R.E. (Drug Abuse Resistance Education; “The History of Dare,” 2019) is perhaps the most well-known example of a popular prevention program which has failed to show positive outcomes in evaluation studies. Specifically, results of several meta-analyses (Ennett, Tobler, Ringwaldt, & Flewwelling, 1994; West & O’Neal, 2004) and evaluation studies (e.g., Becker, Agopian, & Yeh, 1992; Lynam et al., 1999; Thombs, 2000) have shown no differences in rates of adolescent substance use among youth who actively participated in the program compared to those who did not. However, identifying specific mechanisms leading to D.A.R.E.’s ineffectiveness has been challenging due to a lack of consistency in study designs and outcome measures. In general, D.A.R.E.’s emphasis on education about the effects of drugs and about skills needed to avoid initiating substance use is simply insufficient on its own to prevent drug use during adolescence (e.g., Ennett et al. 1994; Thombs, 2000). However, despite these findings, D.A.R.E. continues to be regularly used as a school-based drug prevention program. If further research is unable to identify which aspects of the program are effective and which are not, refinements cannot be

made, resulting in a program that is a waste of money and other resources (e.g., time and training of police officers and school personnel). Thus, evaluating prevention programs is necessary to ensure the best use of scarce community resources.

In summary, despite an ongoing need for evaluation of NPOs and the services provided, many organizations lack the knowledge, capacity, or finances to engage in formal evaluations. However, given the potential for null or iatrogenic effects when programs are not developed according to evidence-based best practices, all NPOs need to undergo evaluation, even those that do not have the funding capability to utilize the evaluation services of large public health agencies. The current study seeks to address these challenges for a small, community-based NPO in the New Orleans, LA area by testing a pilot evaluation procedure and examining preliminary findings from the data collected. Specifically, the current study explored and tested the aims and hypotheses described in the following section.

Aims and Hypotheses

Data for the current study was collected from mother- and youth-reports regarding individual (e.g., delinquency, aggression, and callous-unemotional interpersonal styles) and contextual (e.g., school performance/neighborhood danger/exposure to stressors) characteristics of the boys participating in the SOAS program. These characteristics were examined in the context of boys' level of program involvement (i.e., mentoring and activity participation), as well as the length of time families had been with the organization (i.e., duration). The current study had three primary aims. The first was to obtain and describe information on boys' level of involvement with the program, including time spent participating in 1:1 mentoring and time spent participating in other program activities. The second was to examine boys' problem behavior (i.e., delinquency and aggression) and academic performance (i.e., GPA) in the context of the level of program participation. Finally, the third aim was to examine whether individual (i.e., CU traits) or environmental (i.e., exposure to trauma or stressors) factors moderate the associations between program involvement and antisocial or academic outcomes. To address these aims, the following hypotheses are proposed:

Hypothesis 1. Boys with greater program involvement (i.e., more weekly mentor hours and more activity participation) will report fewer antisocial behaviors (i.e., aggression and delinquency/rule-breaking) and more academic achievement (i.e., higher GPA), compared to boys with less program involvement, both cross-sectionally and longitudinally.

Hypothesis 2. Duration of program involvement will moderate the associations (both cross-sectionally and longitudinally) between program involvement and antisocial/academic outcomes, such that the link between greater program involvement and fewer antisocial behaviors/better academic performance will be stronger for boys who have been in the program

longer. In contrast, boys who have been in the program for a shorter duration will not exhibit significant improvements in behavior or academics with greater program involvement, as the effects of program involvement likely take time to produce measurable change.

Hypothesis 3. Callous-unemotional traits will moderate the association (both cross-sectionally and longitudinally) between program involvement and antisocial behaviors such that boys with high CU traits will exhibit no significant association between program involvement and levels of antisocial behavior. In contrast, boys with low CU traits are expected to show significantly lower antisocial behaviors with greater involvement in program activities.

Hypothesis 4. Exposure to neighborhood danger or environmental stressors will moderate the associations between program involvement and antisocial/academic outcomes both cross-sectionally and longitudinally, such that program involvement will be associated with fewer antisocial behaviors and better academic performance for boys whose mothers report high levels of neighborhood danger or exposure to stressors compared to boys who have been exposed to fewer stressors or neighborhood danger.

Methods

Participants

Son of a Saint (SOAS) is a mentor-based, youth NPO, which targets fatherless boys in the Greater New Orleans area and pairs these boys with male, volunteer mentors. Boys are selected through a referral-based application process. Of the approximately 150 10- to 13-year-old boys who apply for the program, only about 20 boys are selected each year for participation. Once invited into the program, boys begin to develop mentor matches and are required to attend events with their mentors and the program staff. Participation involves completing a minimum of four hours of community service every month, but the actual amount of participation varies across youth. Community service activities vary in type (i.e., service learning, etiquette, mental health, travel, camp) and intensity. Son of a Saint also capitalizes on community activities, such as attending sporting events and local festivals. Boys “graduate” from the program upon entry into college. The dropout rate is remarkably low, with only two boys, to date, terminating the program early. Notably, although SOAS requests that boys participate in at least 2 activities each month and that mentors make weekly contact with mentees, there are no strictly enforced requirements regarding boys’ levels of program involvement. As such, participation in the SOAS program is largely determined by mothers and sons.

Participants in the current sample included 37 mother-reports and 27 youth-reports at T1. Twenty-one mother-reports were collected at T2. The full sample was primarily African-American (81%) and the majority of mothers (60%) held some form of post-secondary degree (e.g., Associate’s or higher, other vocational certification). At T1, boys had been participating at SOAS for an average of 24 months (range = 3-64 months). Primary caregivers were predominantly mothers (n = 35, 94.6%), while the remainder were grandmothers (n = 2, 5.4%).

Boys' ages ranged from 10-18 years old ($M = 13.27$, $SD = 1.99$), and mothers were between 31-62 years old ($M = 40.59$, $SD = 7.85$). Sample characteristics are presented in Table 2.

Procedure and Measures

After obtaining IRB approval, mothers were sent a letter and then recruited via email and asked to participate in the evaluation project supported by SOAS. Mothers viewed a consent video and then read and electronically signed a consent form. Mothers were informed that their participation was voluntary, they could withdraw at any time, and that their participation status would not be shared with SOAS. Moreover, mothers were informed that their participation would take approximately 1 hour and that they would be compensated \$50 for their time. Once mothers completed their survey, they were asked to provide consent for their son to participate. Mothers who chose not to participate themselves had the option of providing consent for their sons. Boys also were asked to provide assent to participate in the study. In the current study, 3 mothers declined to provide consent for sons and 6 were unavailable for boys' survey appointments. Using computers or iPads with links to an electronic survey, boys first watched a short video which described the study and were then asked to provide assent. All boys in the current sample provided assent. Surveys were then completed on the computers or iPads via Qualtrics. Graduate research assistants were on hand to read surveys to boys if needed or to provide assistance completing the survey. Boys were interviewed in groups of 5-10 with at least one graduate student for every 4 boys.

Approximately 10 months after completing the first survey, mothers were approached via email to complete a follow-up assessment of boys' past-year behavior and academic performance. Any mother who completed a survey for her son the previous summer was eligible to participate. The follow-up (T2) survey took approximately 30 minutes to complete and

mothers were compensated \$20 for their time. Mothers were sent emails containing instructions and a link to the Qualtrics survey. A graduate research assistant then contacted mothers who completed the survey to make payment arrangements.

Upon completion of the survey data collection, a graduate student met with staff at SOAS to collect additional information about the program and specific activities in which boys participated. In order to protect participant anonymity, all information was de-identified prior to researchers obtaining access. The research team then cross-checked birthdates of study participants with the list provided by SOAS so that the additional data was only collected for families who had provided consent to participate in the study.

Program involvement measures.

Self-report. To assess boys' level of program involvement, mothers and sons were asked a series of questions at each assessment point regarding a) the number of weekly hours boys spend participating in SOAS activities, and b) the number of weekly hours boys spend with mentors. Only mother-reported Duration values were used for analyses, given that parents are more likely than children to have accurate knowledge of boys' official start dates in the program. However, boys' report of activity participation and mentoring hours were used in models estimating youth-report outcomes, as boys can be reasonably expected to remember how they spend their time at SOAS.

SOAS-Report. Following families' completion of the T2 assessment, SOAS was asked to provide a detailed account of the activities in which each boy participates (e.g., recreational activities, tutoring). The activity records obtained from SOAS included a list of activities offered over the course of the study period (i.e., June, 2018-April 2019) and the birthdates of each boy

who attended the event. Birthdates were cross-checked with demographic information from T1 surveys in order to protect the anonymity of families not participating in the study. The number of activities were summed for each boy, creating a total T2 Activities variable.

Assessment of activities was not initially intended to be measured in different ways at each time point. However, complications arose with regard to the type of activity data that was available from SOAS records and the research team was unable to obtain information on the frequency with which boys participated in group activities between T1 and T2. While not ideal, this complication provided a unique opportunity to capture separate domains of activity participation. Activity participation at T1 was assessed according to the frequency with which boys participated in activities, which is often referred to as the *intensity* of activity participation (Busseri & Rose-Krasnor, 2009; Busseri, Rose-Krasnor, Willoughby, & Chalmers, 2006; Mahoney, Harris, & Eccles, 2006; see also Côté, 1999). In contrast, T2 activities represent the variety of different activities in which boys participated in, known as *breadth* of involvement (e.g., Busseri et al., 2006; Fletcher & Shaw, 2000). While both intensity and breadth of activity involvement have been linked to positive developmental outcomes (e.g., higher academic achievement and reductions in problem behavior; Anderson-Butcher, Newsome, & Ferrari, 2003; Mahoney, Cairns, & Farmer, 2003), evidence suggests that each may contribute to developmental outcomes in different ways. For example, intensity of involvement may be particularly important for identifying and mastering skills and forming lasting prosocial bonds with peers (Busseri et al., 2006; Larson & Verma, 1999). Alternatively, greater breadth of involvement has been linked with fewer risk-taking behaviors and more positive interpersonal functioning (e.g., better relationships with peers and family; Fredricks & Eccles, 2006; Mahoney & Stattin, 2000). As noted by Roth and Brooks-Gunn (2017), few evaluation studies have

assessed activity participation across multiple domains, as most of the extant literature is focused on organized or school-based activities. Thus, although the activity data cannot be compared across time points, the present study will provide important, albeit preliminary, information on both the intensity and breadth of activity participation for boys in the current sample.

Table 1. *Reliability estimates for measures and subscales of main study variables*

Measure/Construct	Youth-Report α	T1 Mother-Report α	T2 Mother-Report α
STAB-Physical Agg.	.90	.84	.88
STAB-Rule-Breaking	.70	.54	.60
STAB-Tot. Delinquency	.91	.82	.82
PRA-Reactive Agg.	.68	.78	.80
PRA-Proactive Agg.	.78	.63	.88
PRA-Total Aggression	.77	.73	.70
ICU- Total CU Traits	.82	.86	.87

Antisocial behavior measures.

Subtypes of Antisocial Behavior Questionnaire (STAB; Burt & Donnellan, 2009). The STAB is a 38-item questionnaire designed to measure the frequency of youth’s engagement in distinct types of antisocial behavior. Each item is rated on a 5-point scale (1 = *Never*, 2 = *Hardly ever*, 3 = *Sometimes*, 4 = *Frequently*, and 5 = *Nearly all the time*). Total subscale scores are obtained by computing the mean for each response set. The Rule-Breaking subscale consists of 13 items (e.g., “Stole property from school or work”), the Physical Aggression subscale consists of 14 items (e.g., “Felt like hitting people”), and the Social Aggression subscale consists of 11 items (e.g., “Made fun of someone behind their back”). In the current study, only the Physical Aggression and Rule-Breaking subscales are included in analyses. In order to reduce the number of parameters used in study analyses and to minimize multicollinearity, Physical Aggression and Rule-Breaking were combined into a single “Delinquency” construct. Bivariate correlations between the subscales indicated that the Delinquency construct could be used reliably across

informants and assessment points. That is, Physical Aggression and Rule-Breaking were highly correlated (i.e., Pearson's $r \geq .30$) among T1 mother-report ($r = .47$), youth-report ($r = .74$) and T2 mother-report ($r = .32$; see Tables 4 and 5).

Mother-report. The STAB has been validated for use as a parent-report tool and has demonstrated good internal consistency in prior studies (Cronbach's $\alpha = .89$; van der Veen-Mulders, Nauta, Timmerman, van den Hoofdakker, & Hoekstra, 2017). In the present analyses, numerous items were removed at both T1 and T2 due to zero variance (i.e., no participant endorsement of the item). Items were removed as follows: Physical aggression- #32) *Attacked someone with a weapon* (T2), #33) *Used a weapon or force to get something he wants* (T2), #34) *Played with a weapon just for fun* (T1 and T2), and #36) *Harmed an animal or person just for fun* (T1 and T2). Rule-breaking- #2) *Broke into a store, mall, or warehouse* (T2), #7) *Shoplifted things* (T2), #13) *Stole a bicycle* (T2), #16) *Stole property from school* (T2), #19) *Left home without telling* (T1), #25) *Suspended from school* (T1), #40) *Used cigarettes* (T1), and #42) *Used other illicit drugs* (T1). In the current sample, reliability of the physical aggression subscale was good, while the reliability of the rule-breaking subscale was fair. Reliability for the overall "Delinquency" construct was good. Cronbach's α values for all subscales and study constructs across assessment point and informant are presented in Table 1.

Youth-report. Both Physical Aggression and Rule-Breaking subscales have demonstrated good internal consistency in validation studies using community youth samples (e.g., Cronbach's $\alpha = .85$ for Physical Aggression, $\alpha = .87$ for Rule-Breaking; Burt & Donnellan, 2009). The STAB questionnaire has demonstrated good convergent validity with similar measures of antisocial behavior. In a validation study conducted by the STAB authors, Physical Aggression was positively and significantly correlated ($r = .67$) with the Physical Aggression subscale of the

Externalizing Spectrum Model (ESM; Krueger et al., 2007), even when the effects of other subscales were statistically controlled for. Similarly, the Rule-Breaking subscale was strongly correlated with ESM measures of Theft ($r = .59$, controlling for other subscales), Rebelliousness ($r = .46$, controlling for other subscales), and Destructive Aggression ($r = .53$, controlling for other subscales). In the current study, the physical aggression subscale demonstrated excellent reliability, while rule-breaking was adequate. Overall reliability for youth-reported physical aggression and rule-breaking (i.e., “Delinquency”) was excellent (see Table 1). Three items were not endorsed by any participants and thus were eliminated from reliability analysis. Items removed included one physical aggression question (#33. *Used a weapon or force to get what I want*) and two items from the rule-breaking subscale (#2. *Broke into a store, mall, or warehouse*; #33. *Failed to pay debts*).

Proactive and Reactive Aggression Questionnaire (PRA; Dodge & Coie, 1987). The PRA is a 6-item questionnaire designed to measure proactive and reactive types of aggressive behavior. The measure includes 3 items assessing reactive aggression (e.g., “When I have been teased or threatened, I get angry easily and strike back”) and 3 items assessing proactive aggression (e.g., “I get other kids to gang up on somebody that I do not like”). Items are rated on a 5-point scale (1 = *Never*, 2 = *Very Rarely*, 3 = *Sometimes*, 4 = *Often*, and 5 = *Almost always*). Subscale scores are obtained by summing item responses. A total aggression score can be obtained by calculating the mean of the 2 subscale scores. Although reactive and proactive aggression are typically highly correlated (e.g., Price & Dodge, 1989; see also Crapanzano, Frick, & Terranova, 2010), the factor structure of the PRA has consistently indicated that the 6 items represent distinct dimensions of aggressive behavior (e.g., Dodge & Coie, 1987; Poulin & Boivin, 2000). Independent validation studies have indicated excellent reliability for both

proactive and reactive aggression subscales (e.g., Cronbach's $\alpha = .92$ and $.90$, respectively; Poulin & Boivin, 2000).

The reactive and proactive aggression subscales were significantly correlated across informants and assessment points and were therefore combined to form a single "Aggression" variable (see Tables 4 and 5). In the current sample, reliability on both the reactive and proactive aggression subscales demonstrated a range of reliability, from fair to acceptable to good. Reliability for the full PRA scale (i.e., proactive and reactive aggression) was acceptable across groups and assessment points (see Table 1). One item on the proactive aggression subscale was not endorsed by any participant at T2 and was removed (#4. *My son gets other kids to gang up on somebody that he does not like*).

Academic Performance

Mothers and sons each reported grades in Math, English/Language Arts, Science, and Social Studies/History. Mothers reported grades again at T2. Each set of responses were averaged to obtain boys' total GPA.

Contextual moderating variables

Inventory of Callous-Unemotional Traits (ICU; Frick, 2004).

Mother-Report. The parent-report version of the ICU (ICU-PR; Frick, 2004), is a 24-item self-report questionnaire designed to assess parent report of CU traits in youth. Items are scored on a 4-point scale (0 = *not at all true*, 1 = *somewhat true*, 2 = *very true*, and 3 = *definitely true*). The measure consists of twelve positively-worded items (e.g., "My child expresses his feelings openly") that are reverse scored such that 0 = *definitely true*, 1 = *very true*, 2 = *somewhat true*, and 3 = *not at all true*, and twelve negatively-worded items (e.g., "The feelings of others are

unimportant to him”). Prior research has demonstrated good internal consistency (e.g., Cronbach’s $\alpha = .85$; Hawes et al., 2014). Additionally, White, Cruise, and Frick (2009) reported that the ICU-PR significantly predicted risk of antisocial behaviors and did so more reliably than the youth self-report version of the measure. For the current sample, overall reliability was good (Table 1).

Youth-Report. Boys were administered a shortened version of the ICU measure, the ICU-10 (ICU-10; Frick, 2004; Ray, Frick, Thornton, Steinberg, & Cauffman, 2016). Like the parent-report version, items are scored on a 4-point scale (0 = *Not at all true*, 1 = *Somewhat true*, 2 = *Very true*, and 3 = *Definitely true*). The ICU-10 contains 10 items from the original 24-item Youth Self-Report and includes 8 positively worded items (e.g., “I work hard on everything I do”) and 2 negatively worded items (e.g., “I do not care about doing things well”). Positively worded items are reverse scored, such that 0 = *Definitely true* and 3 = *Not at all true*. Items are summed to create a total CU score. The internal consistency of the ICU-10 has not been found to significantly differ from the full ICU on measures of internal consistency (Cronbach’s $\alpha = .78$ for both ICU-10 and full ICU when administered to participants in the same sample; Ray et al., 2015). The measure has also demonstrated good criterion validity, showing positive correlations with proactive aggression ($r = .25, p < .001$), school misconduct ($r = .32, p < .001$), delinquency ($r = .30, p < .001$), and reactive aggression ($r = .27, p < .001$; Ray et al., 2015). In the current sample, reliability was good (Table 1).

Neighborhood danger (adapted from the Me and My Neighborhood Questionnaire; Pittsburg Youth Study, 1991). Mothers completed a 20-item measure assessing dangerous or risky events that have occurred in their neighborhood during the past year. Eight items assessed direct experience with neighborhood danger (e.g., “During the past year, how often did you see a

gang fight occur near your home?”), 4 items assessed awareness of neighborhood danger or risk (e.g., “During the past year, how often did you hear neighbors complaining about crime in your neighborhood?”), and 8 items assessed dangerous events experienced by close family or friends (e.g., “During the past year, how often did a family member get robbed or mugged?”). Items are rated on a 4-point scale (0 = *never*, 1 = *once*, 2 = *a few times*, and 3 = *a lot*). However, some events represent more severe danger than others, which may artificially inflate overall scores (e.g., witnessing a shooting near your home one time may represent a higher degree of danger than frequently hearing people in your neighborhood complain about crime). Therefore, in the current study, scores were dichotomized, such that 0 = event did not occur in the past year and 1 = event occurred at least once in the past year. Responses were then summed to create a total score in which higher numbers reflect a greater number of distinct events occurring over the past year. This procedure has been used in prior studies as a means to control for variability in the severity of dangerous events (e.g., Sapotichne, 2015).

Childhood Trust Events Survey (CTES; Boat, Baker, & Abrahamson, 1996). The CTES is a 26-item index used to assess children’s exposure to traumatic or stressful events during their lifetime. In order to comply with IRB and SOAS requirements, all items related to physical/domestic violence and sexual abuse were removed from the measure before the index was added to the survey. The remaining 16 items are designed to screen for traumatic stressors related to witnessing (e.g., “has your child ever seen a friend killed?”) or being victimized by others (e.g., “has someone ever robbed or tried to rob your child or your family with a weapon?”), as well as items related to health (e.g., “was your child ever in a really bad accident, such as a serious car accident?”) and exposure to natural disasters (e.g., “was your son ever in a disaster such as a tornado, hurricane, fire, big earthquake, or flood?”). Response options are

dichotomized into Yes/No choices and summed to obtain a score reflecting the number of distinct traumatic stressors boys have experienced during their lifetimes. Given that the CTES is used as a simple screening index for self-reported traumatic events, validity and reliability estimates have not been ascertained in the literature (Pearl et al., 2012). Scores on the CTES were correlated with scores on the Neighborhood Danger index across informants and times points (see Tables 4 and 5), and were therefore combined to create a single “Trauma/Stressors” composite score.

Table 2. *Demographic characteristics of sample*

Characteristic	<i>n</i>	%
Race (<i>Mother</i>)		
Black/African American	30	81.1
White/Caucasian	6	16.2
Other	1	2.7
Race (<i>Son</i>)		
Black/African American	22	59.5
White/Caucasian	2	5.4
Other	3	8.1
Ethnicity (<i>Mother</i>)		
Hispanic/Latino	2	5.4
Not Hispanic/Latino	35	94.6
Ethnicity (<i>Son</i>)		
Hispanic/Latino	2	8.3
Not Hispanic/Latino	22	91.7
Education Level (<i>Mother</i>)		
< High School	2	5.4
High School/GED	5	13.5
Some college	9	24.3
Associates	5	13.5
Bachelors	6	16.2
Masters or higher	8	21.6
Other vocational training	2	5.4
Employment Status (<i>Mother</i>)		
Employed full-time	17	45.9
Currently unemployed	1	2.7
Stay-at-home mom	2	5.4
Income (<i>Mother</i>)		
< \$20,000	7	18.9
\$20,000-\$39,999	2	5.4
\$40,000-\$59,999	8	21.6
\$60,000-\$79,999	2	5.4
\$80,000+	1	2.7

Table 3. Group means, standard deviations, and range on indicators of main study constructs

	Boys N = 27		Mothers (T1) N = 37		Boys/ T1 Mothers Comparison	Mothers (T2) N = 21		T1/ T2 Mothers Comparison
	M (SD)	Range	M (SD)	Range	t (df)	M (SD)	Range	t (df)
Delinquency								
PhyAgg	24.78 (9.08)	11-47	20.91 (5.40)	14-35	-1.96 (39.87) [†]	20.10 (6.61)	13-37	.51 (54)
RulBrk	13.59 (4.01)	5-22	12.20 (1.49)	11-17	-1.72 (31.56) [†]	12.85 (2.89)	11-21	-.97 (26.52)
TotDel	38.37 (12.40)	16-69	33.11 (6.25)	25-51	-2.01 (36.13) [*]	32.95 (8.05)	24-55	.08 (54)
Aggression								
ProAgg	3.89 (1.65)	3-9	3.31 (0.99)	2-7	-1.60 (40.19)	3.14 (0.65)	2-9	.70 (54)
ReactAgg	7.15 (2.57)	3-12	6.03 (2.42)	3-11	-1.76 (60) [†]	7.05 (3.29)	3-13	-1.33 (54)
TotPRA	5.52 (1.86)	3-10	9.34 (2.91)	6-16	6.28 (58.23) ^{**}	10.19 (3.44)	6-16	-.99 (54)
Academics								
GPA	2.70 (0.67)	1.50-4.00	2.68 (0.74)	1.25-4.00	-.07 (62)	2.35 (0.81)	0.75-3.50	1.57 (55)
Prg. Involve								
1:1 Ment	3.54 (2.50)	0-10	2.88 (2.27)	0-8	-1.05 (57)	4.95 (4.26)	0-16	-2.05 (27.19) [*]
Activities ^A	4.00 (1.89)	2-7	3.53 (2.09)	0-8	-.89 (57)	12.89 (6.92)	2-27	-.00 (43.06)
Moderators								
Duration	24.24 (26.41)	1-96	23.82 (18.72)	3-64	---	---	---	---
CU Traits ^B	8.85 (5.76)	0-18	24.44 (9.99)	6-48	1.74 (61) [†]	24.24 (10.01)	7-44	.08 (54)
ND ^C	---	---	6.00 (4.51)	0-17	---	---	---	---
Ch.Stress ^C	---	---	3.41 (1.83)	0-8	---	---	---	---
Tot. Stress^C	---	---	4.70 (2.79)	1.5-11.5	---	---	---	---

Note: [†] $p < .10$; ^{*} $p < .05$; ^{**} $p < .01$ PhyAgg = Physical Aggression; RulBrk = Rule-Breaking; ReactAg = Reactive Aggression; ProAgg = Proactive Aggression; Prg. Involve = Program Involvement; ND = Neighborhood/Community Danger; Ch. Stress = Childhood Stressful Events; TotDel = Total Delinquency composite; TotPRA = Total Aggression composite Tot Stress = Neighborhood Danger/Ch.Stress composite

^A Activity participation at T2 assessed according to number of distinct activities boys participated in, T1 assessed according to average weekly hours; actual mean values of activity participation shown, t-tests conducted on standardized activity variables due to difference in measurement across time points

^B Child report CU traits assessed using abbreviated (10-item) version of the ICU. Mothers' report uses full 24-item version. Mother-report CU trait scores were recalculated using only those items included on the ICU-10 for t-test comparison with youth-report

^C Neighborhood danger and childhood stress events assessed by mothers at T1 only

Table 4. *Bivariate correlations of main study constructs at T1*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Son Age	*	.50**	-.36*	-.07	-.16	-.32*	-.24 ^t	-.34*	-.20	.04	-.14	.17	-.33*	-.24 ^t	-.05
2. Duration	.60**	*	-.18	-.01	-.19	.20	.20	-.21	-.25 ^t	.02	-.09	.11	-.15	-.15	-.04
3. Phy. Agg	-.05	-.13	.18	.47**	.69**	.50**	-.07	.06	-.04	.47**	.24 ^t	.11	.98**	.74**	.23 ^t
4. RulBrk	.10	-.11	.74**	.19	.39*	.49**	-.01	-.00	-.06	.31*	-.10	-.03	.64**	.49**	-.09
5. ReactAg	.08	-.21	.59**	.60**	.38*	.34*	-.23 ^t	.07	-.10	.43**	.26 ^t	.15	.69**	.95**	.26 ^t
6. ProAg	-.05	-.20	.49**	.60**	.60**	-.34*	.38*	-.05	.02	.20	.20	.21	.55**	.62**	.23 ^t
7. GPA	-.24	.04	-.42*	-.36*	-.44*	-.20	.83**	.09	.14	-.30*	-.18	-.10	-.06	-.06	-.18
8. Mentoring	-.31 ^t	-.34*	.21	.21	.31 ^t	.25	.10	.32 ^t	.74**	-.39*	.20	-.03	.05	.04	.16
9. Activities	-.17	-.29 ^t	.33*	.44*	.22	.34*	-.03	.61**	.16	-.53**	.18	.02	-.05	-.08	.16
10. CU Traits	.21	-.08	.52**	.50**	.51**	.53**	-.61**	.23	.19	.15	.13	.05	.48**	.42**	.12
11. ND	-.08	-.21	.21	.05	.14	.19	-.15	.15	.23	.19	*	.40**	.19	.29*	.96**
12. Ch.Stress	.17	.02	.45*	.21	.23	.20	-.30 ^t	.18	.16	.18	.55**	*	.09	.19	.65**
13. TotDelinq	-.01	-.13	.98**	.86**	.63**	.55**	-.43*	.22	.38*	.54**	.17	.40*	*	.76**	.18
14. Total PRA	.03	-.23	.61**	.66**	.94**	.84**	-.39*	.32 ^t	.30 ^t	.58**	.18	.24	.67**	*	.30*
15. Tot. Stress	-.02	-.17	.29 ^t	.10	.18	.21	-.20	.17	.23	.21	.98**	.72**	.25	.21	*

Note: Lower diagonal = Boys' report

t = p < .10; *p < .05; **p < .01

Upper diagonal = T1 mothers' report

Middle diagonal = mother-son agreement (when applicable)

PhyAgg = Physical Aggression; RulBrk = Rule-Breaking; ReactAg = Reactive Aggression; ProAg = Proactive Aggression; ND = Neighborhood/Community Danger; Ch. Stress = Childhood Stressful Events; TotDelinq = Total Delinquency composite; Total PRA = Total Aggression composite Tot. Stress = ND/Ch. Stress composite

Table 5. *Bivariate correlations of study constructs at Time 2*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Son Age	*														
2. Duration	.57**	*													
3. Phy. Agg	-.66**	-.41*	*												
4. RulBrk	.03	-.01	.32^t	*											
5. ReactAg	-.38*	-.42*	.67**	.36^t	*										
6. ProAg	-.43*	-.21	.59**	.41*	.31^t	*									
7. GPA	.11	.16	-.08	-.31^t	-.13	.04	*								
8. Mentoring	-.31^t	-.35^t	.11	-.07	.23	-.22	-.37*	*							
9. Activities	-.16	-.12	-.06	-.34^t	-.19	.06	.12	.20	*						
10. CU Traits	-.01	-.07	.20	-.02	.32^t	-.14	-.29	-.09	-.38*	*					
11. ND	-.32^t	-.10	.31^t	-.08	-.13	.26	-.12	.10	.24	-.17	*				
12. Ch.Stress	.02	.15	-.01	-.21	-.10	-.02	-.12	.24	.30^t	-.04	.39*	*			
13. TotDelinq	-.53**	-.34^t	.94**	.62**	.68**	.63**	-.18	.07	-.17	.16	.23	-.08	*		
14. Total PRA	-.44*	-.45*	.75**	.42*	.98**	.32^t	-.12	.17	-.17	.28	-.08	-.10	.77**	*	
15. Tot. Stress	-.26	-.05	.26	-.13	-.14	.22	-.14	.16	.29	-.15	.97**	.62**	.17	-.09	*

Note: t = $p < .10$; * $p < .05$; ** $p < .01$

PhyAgg = Physical Aggression; RulBrk = Rule-Breaking; ReactAg = Reactive Aggression; ProAg = Proactive Aggression; ND = Neighborhood/Community Danger; Ch. Stress = Childhood Stressful Events; TotDelinq = Total Delinquency composite; Total PRA = Total Aggression composite

Data Analysis Plan

Preliminary data analyses. All study variables were first examined to ensure that assumptions of normality had not been violated. Descriptive analyses were then conducted regarding the nature of the sample. These data included ages of mothers and sons, race/ethnicity, mother's income/employment status/education level. Next, bivariate correlations, means, and variability were computed for each set of indicator variables that were used to create study constructs. Since antisocial behavior often increases as boys progress through adolescence (Broidy et al., 2003), age was examined as a potential covariate and was controlled for in any model in which an outcome varied significantly as an effect of boys' ages. Independent samples t-tests were then used to determine whether mothers and sons provided significantly different responses on main study variables, as well as whether mothers' responses differed significantly between T1 and T2.

Main study hypotheses. To test Hypothesis 1, that boys with greater SOAS involvement will report fewer antisocial behaviors and better academic performance compared to boys with less program involvement, multiple regression modeling was conducted using Mplus (Muthén & Muthén, 2012). Mother- and son-reported outcomes were tested separately. Mentoring hours and activity involvement were entered individually as predictors to determine whether specific types of program involvement differentially affect levels of academic performance (i.e., GPA) or antisocial behavior (i.e., delinquency and aggression). These procedures were then replicated using data collected at T2 to assess cross-sectional effects at the 1-year follow-up. Finally, the T2 models were tested with each respective T1 outcome variable (i.e., Delinquency, Aggression, and GPA) included as a covariate to evaluate whether program involvement predicted antisocial and academic outcomes at T2 after controlling for initial levels.

To test hypotheses 2-4, regression equations were computed to determine the extent to which each moderator (i.e., Duration, CU traits, Stressors/Trauma) interacted with each aspect of program involvement to predict antisocial/academic outcomes. Control variables (when applicable) and main effects (e.g., of Mentoring and Activity Intensity or Breadth) were first entered individually into the regression equations in order to assess the main effects of each aspect of program involvement. Next, moderators were entered into the equation to determine whether main effects of hypothesized moderators would emerge. Finally, interaction terms were created for each moderator and each aspect of program involvement and added to the regression. For example, to test Hypothesis 2, Duration (mean-centered), mentoring hours (mean-centered), and activity participation (mean-centered) were entered into the first set of regression equations, followed by a Duration x Mentoring interaction term, then retested using a Duration x Activities interaction term. Post-hoc probing was conducted for any significant interactions to determine the strength and direction of effects. As with Hypothesis 1, procedures were replicated using T2 data to assess additional cross-sectional effects, as well as longitudinal effects controlling for T1 outcomes.

Results

Aim 1: Description of Program Involvement

At the time of the first assessment, boys reported spending an average of 3.54 hours per week in 1:1 mentoring sessions (range = 0-10 hours) and an average of 4 hours per week participating in other (i.e., non-mentored) activities (range = 2-7 hours). Mothers reported slightly fewer weekly 1:1 mentoring hours ($M = 2.88$, range = 0-8) and weekly activity hours ($M = 3.53$, range = 0-8) compared to boys' report, although these differences were not significant. At T2, mothers reported significantly more weekly 1:1 mentoring hours ($M = 4.95$, range = 0-16) compared to T1 assessment. This increase is likely due to the increased number of mentors volunteering at SOAS between T1 and T2 (Program Director, personal communication, May 9, 2019).

Review of activity records indicated that, on average, boys participated in 14 activities between T1 and T2 (range = 2-27). This number does not account for activities in which boys participated on multiple occasions and is likely somewhat lower than the actual number of past-year activities. The majority of boys (54%) participated in 10 or more different activities during the study period. All boys attended at least one Etiquette Meal, during which boys learn proper table manners while dining at upscale restaurants throughout the community. Additionally, all boys attended the required annual SOAS fundraising gala. Recreational (e.g., attending professional sporting events, horseback riding, etc.) and Life Skills (e.g., Etiquette Meals, Financial Literacy, Public Speaking and Attire, etc.) were offered more frequently than Community Service, Vocational, or Academic activities, and were therefore the most commonly attended types of activity for all boys.

Aim 2: Associations between Program Involvement and Antisocial/Academic Outcomes

Preliminary analyses. All study constructs were first examined to determine whether the distributions met assumptions of normality. All skewness and kurtosis values were within the acceptable range (i.e., less than 3.00), indicating that the constructs were normally distributed. Table 3 presents a descriptive summary of all main study variables. Group differences were examined on all study constructs between mothers who did and did not participate at T2 in order to determine the extent to which T2 participation was random. Independent samples t-tests did not reveal any significant differences between “T1-only” and “T1/T2” mothers on any variable, indicating that no selection bias had occurred across assessment points.

Bivariate correlations were examined to test associations between program involvement, antisocial/academic outcomes, and contextual factors in the current sample. Table 4 contains correlations among all main study variables for mothers and sons at T1. Table 5 presents the same associations for mothers at T2. Among mothers at T1, age was significantly associated with nearly all outcome variables (see Table 4, upper diagonal). Thus, boys’ age was controlled for in all mother-report models. Since age was uncorrelated with all youth-reported outcome variables (Table 4, lower diagonal), however, age was *not* controlled when estimating youth-report models.

Hypothesis 1. Hypothesis 1 stated that more program involvement (measured by breadth and intensity of activity participation and mentoring hours) would be associated with lower levels of antisocial behavior and better academic performance than less program involvement.

Time 1. At the bivariate level, among youth-report, participation in 1:1 mentoring was marginally and positively related to total aggression scores ($r = .32$; $p = .06$), indicating that boys

who exhibited more aggressive behavior participated in somewhat more mentoring hours than boys who were less aggressive (Table 4, lower diagonal). T1 Mentoring was not significantly related to any other study constructs in the boys' sample. Activity Intensity was significantly and positively related to boys' self-reported total delinquency ($r = .38; p < .05$) and was marginally positively related to total aggression ($r = .30; p = .08$) indicating that boys with higher levels of overall antisocial behavior participated in SOAS-related activities somewhat more frequently than boys with lower levels of antisocial behavior. Activity Intensity was not associated with any other study construct in the boys' sample. Mentoring and Activity Intensity were unrelated to all other study constructs at T1 mother-report (Table 4, upper diagonal). Contrary to expectations, none of the regression models revealed significant overall effects, and no significant main effects of either mentoring or activity intensity were detected among mothers or sons (Table 6).

Time 2. Bivariate cross-sectional analyses at T2 indicated that mentoring was negatively and significantly related to GPA ($r = -.37; p < .05$), such that boys with higher GPAs participated in fewer mentoring hours. No other study constructs were related to mentoring at T2. Results of cross-sectional regression analyses are presented in Table 7 and revealed a trend toward a significant main effect of breadth of activity participation on Delinquency when age and Mentoring were included as covariates in the model ($\beta = -.30, p = .07$). This finding suggested that boys who participated in more activities displayed marginally fewer mother-reported delinquent behaviors than boys who participated in fewer activities. A significant and negative main effect of mentoring on T2 GPA emerged ($\beta = -.40, p < .05$), indicating that more mentoring hours were associated with lower T2 GPA. This finding is in contrast to the hypothesis that more 1:1 mentoring hours would be associated with better grades in school. When T1 outcomes were added to the models as covariates, the only significant main effects in each model were T1

Delinquency ($\beta = .79, p < .01$), T1 Aggression ($\beta = .50, p < .05$), and T1 GPA ($\beta = .79, p < .01$), suggesting that prior scores on outcome variables predicted the majority of the variance in each model.

Hypothesis 2. Hypothesis 2 stated that the duration of boys' program involvement would moderate the associations between activity participation, mentoring, and antisocial/academic outcomes. Specifically, participation in mentoring or activities was expected to be more strongly associated with fewer antisocial behaviors and better academic performance among boys who have been attending SOAS for a longer period of time compared to boys who have attended for shorter durations. Regression results for T1 analyses are summarized in Table 6 and T2 are presented in Table 7.

Time 1. Contrary to expectations, no significant main or interactive effects of Duration were detected among mothers or sons on the Delinquency construct. A significant Duration x Activity Intensity interaction emerged on the Aggression construct among mothers' report at T1 ($\beta = .33, p < .05$; Table 7, Panel A). Post hoc probing of the interaction was used to determine the conditional effects of the intensity of activity participation on Aggression for boys with shorter and longer duration of involvement. Examination of the simple slopes at low, medium, and high duration indicated that for boys who were in the program for a shorter duration, participating more frequently in program activities was associated with lower levels of aggression ($\beta = -.61, p < .05$). In contrast, no significant association emerged between intensity of activity participation and aggression for boys who were in the program for a longer duration ($\beta = .54, p = .29$). This finding indicates that the relationship between lower intensity of activity participation and aggressive behavior may be stronger for boys who are new to the program. Examination of regions of significance indicated that the interaction became significant at 11

months, indicating that less frequent activity participation is associated with higher levels of aggression for boys who have been in the program for less than 11 months. Conversely, there was not a significant association between lower intensity of activity participation and aggressive behavior among boys who had been in the program for longer than 11 months. The interaction is depicted in Figure 1.

Duration was a significant predictor of GPA ($\beta = .48; p < .01$) when accounting for the effects of mentoring or intensity of activity participation among mothers' report, indicating that higher GPA may be associated with longer duration of involvement at SOAS. More extensive longitudinal analyses are required to fully examine this effect. The interaction terms for Duration x Mentoring and Duration x Activity Intensity were not significant (Table 6, Panel A). No other effects approached significance (see Table 6).

Time 2. Regression results for T2 models are presented in Table 8. Cross-sectional analyses revealed a trend towards a significant effect of T2 breadth of activity participation on T2 Delinquency ($\beta = -.30, p = .09$). This finding partially supports the study hypotheses, in that more participation in program activities was associated with lower levels of delinquency at T2. Breadth of activity participation was significant beyond the effect of duration of program involvement and mentoring hours. Similarly, T2 Mentoring was marginally and negatively related to T2 GPA ($\beta = -.40, p = .08$) beyond the effects of Duration, indicating that boys who are struggling academically spend more time with mentors, irrespective of the duration of involvement. Neither the Duration x T2 Mentoring nor Duration x T2 Activity Breadth interaction terms were significant when added to any other models. When T1 outcomes were added to each model to test for longitudinal effects, the main effects of Activity Breadth on T2 Delinquency/Aggression and the main effect of Mentoring on T2 GPA were no longer

significant. As in Hypothesis 1, T1 Delinquency, Aggression, and GPA were the only significant predictors of the respective T2 outcomes.

Table 6. Hierarchical regression analyses estimating main and interactive effects of program involvement on antisocial and academic outcomes at T1 (Hyp 1 and 2).

	<i>Panel A: T1 Mothers</i>					<i>Panel B: T1 Sons</i>				
Variable	R ²	β	SE	t	p	R ²	β	SE	t	p
	T1 Delinquency					T1 Delinquency				
Age		-.35	.22	-1.61	.11		--	--	--	--
Mentoring	.13	.04	.29	.15	.88	.15	-.02	.29	-.05	.96
Activities		-.15	.21	-.72	.47		.39	.34	1.15	.25
Duration	.13	-.01	.16	-.08	.94	.15	-.03	.15	-.17	.86
Dur x Ment	.13	.00	.17	.02	.98	.15	-.10	.32	-.31	.75
Dur x Act	.14	.11	.15	.76	.45	.19	-.33	.27	-1.23	.22
	T1 Aggression					T1 Aggression				
Age		-.24	.20	-1.21	.23		--	--	--	--
Mentoring	.08	.12	.26	.47	.64	.12	.23	.40	.57	.57
Activities		-.22	.23	-.93	.35		.16	.40	.40	.69
Duration	.06	-.06	.20	-.30	.77	.13	-.12	.19	-.64	.52
Dur x Ment	.11	.18	.17	1.05	.30	.13	-.04	.44	-.08	.94
Dur x Act	.18	.33	.15	2.16	< .05	.18	-.33	.29	-1.13	.30
	T1 GPA					T1 GPA				
Age		-.25	.17	-1.49	.14		--	--	--	--
Mentoring	.07	-.13	.31	-.41	.68	.02	.19	.27	.73	.47
Activities		.19	.28	.65	.51		-.15	.27	-.55	.59
Duration	.24*	.48	.14	3.42	< .01	.03	.08	.26	.29	.77
Dur x Ment	.24*	.03	.21	.14	.89	.04	-.13	.56	-.23	.82
Dur x Act	.26*	.16	.20	.80	.42	.16	.58	.36	1.62	.11

Note: ^t $p < .10$; * $p < .05$; ** $p < .01$

Figure 1. *Duration x Activity Participation predicting total proactive/reactive aggression – mothers' report (Hyp. 2)*

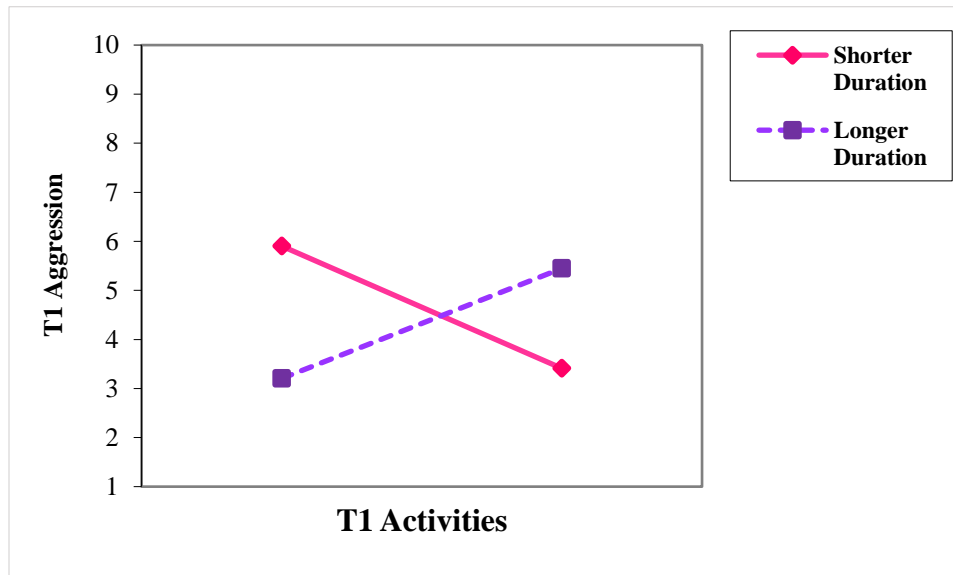


Table 7. *Hierarchical regression analyses estimating cross-sectional and longitudinal effects of program involvement on antisocial and academic outcomes at T2 follow-up (Hyp. 1, T2)*

Variable	R ²	β	SE	t	p
T2 Delinquency					
Age		-.54	.26	-2.10	< .05
T2 Mentoring	.32 ^t	-.01	.24	-.05	.96
T2 Activities		-.30	.17	-1.82	< .10
T1 Delinquency	.76**	.79	.13	6.26	< .01
T2 Aggression					
Age		-.41	.21	-1.97	< .05
T2 Mentoring	.25 ^t	.13	.19	.70	.49
T2 Activities		-.31	.17	-1.79	< .10
T1 Aggression	.40*	.50	.23	2.19	< .05
T2 GPA					
Age		.02	.26	.06	.95
T2 Mentoring	.17	-.40	.21	-1.97	< .05
T2 Activities		.20	.19	1.01	.31
T1 GPA	.50**	.77	.23	3.39	< .01

Note: ^t $p < .10$; * $p < .05$; ** $p < .01$; all T2 main effects became non-significant after adding T1 outcomes into the models

Table 8. Hierarchical regression analyses estimating cross-sectional and longitudinal main and interactive effects of program involvement and duration on antisocial and academic outcomes at T2 (Hyp. 2, T2)

Variable	R ²	β	SE	t	p
T2 Delinquency					
Age		-.52	.33	-1.59	.11
T2 Mentoring		-.02	.26	-.07	.95
T2 Activities	.32^t	-.30	.18	-1.71	< .10
Duration		-.03	.27	-.12	.90
Dur x T2 Ment	.33*	.11	.37	.30	.76
Dur x T2 Act	.32^t	-.08	.29	-.28	.78
T1 Delinquency	.76**	.79	.16	4.86	< .01
T2 Aggression					
Age		-.29	.30	-.97	.33
T2 Mentoring		.09	.21	.44	.66
T2 Activities	.28^t	-.31	.20	-1.58	.11
Duration		-.23	.28	-.84	.40
Dur x T2 Ment	.31*	.18	.33	.56	.58
Dur x T2 Act	.28^t	.01	.28	.03	.98
T1 Aggression	.42*	.48	.27	1.78	< .10
T2 GPA					
Age		-.01	.36	-.02	.98
T2 Mentoring	.18	-.40	.23	-1.73	< .10
T2 Activities		.20	.21	.95	.34
Duration		.05	.42	.11	.91
Dur x T2 Ment	.34*	.45	.38	1.21	.23
Dur x T2 Act	.18	-.02	.48	.12	.91
T1 GPA	.51*	.79	.21	3.80	< .01

Note: ^t $p < .10$; * $p < .05$; ** $p < .01$; all T2 main and interaction effects became non-significant after adding T1 outcomes into the models

Aim 3: The Role of Contextual Factors

Hypothesis 3. Hypothesis 3 stated that callous-unemotional (CU) traits would moderate associations between program involvement, delinquency, and aggressive behavior, such that boys with low CU traits would exhibit significantly fewer delinquent or aggressive behaviors associated with more program involvement, while boys with high CU traits would not. Results of T1 regression analyses are summarized in Table 9. Table 10 presents regression results at T2.

Time 1. Mother-report at T1 revealed significant and negative bivariate associations between mentoring hours and CU traits ($r = -.39; p < .05$) and between intensity of activity participation and CU traits ($r = -.53; p < .01$), indicating that boys with higher CU traits spend less time participating in mentoring and activities than boys with low CU traits (see Table 4, upper diagonal). Significant and positive main effects of CU traits on mother-reported Delinquency ($\beta = .60, p < .01$) and Aggression ($\beta = .51, p < .01$), and youth-reported Delinquency ($\beta = .50, p < .01$) and Aggression ($\beta = .53, p < .01$), emerged. No significant main effects of mentoring or activities or interactive effects of CU x Activity Intensity or CU x Mentoring interactions were found for mother- or youth-report models. Regression results are presented in Table 9.

Time 2. T2 correlational analyses indicated that breadth of activity participation at T2 was negatively and significantly related to CU traits ($r = -.38; p < .05$; Table 5). Results of T2 regression analyses are presented in Table 10. Cross-sectional analyses of data collected at T2 revealed no significant main effects of CU traits, Mentoring, or Activity Breadth in models predicting T2 Delinquency or T2 Aggression. Further, no significant interactive effects were found for CU x Mentoring or CU x Activity Breadth in either model suggesting that levels of CU traits do not influence the effects of program involvement on antisocial outcomes at T2. When

T1 Delinquency/Aggression scores were added to the models to test for longitudinal effects, T1 Delinquency significantly predicted T2 Delinquency beyond the effects of T2 Mentoring, Activity Breadth, and CU traits ($\beta = 1.08$; $p < .01$) Aggression scores at T1 were not significant predictors of Aggression at T2.

Table 9. *Regression analyses estimating main and interactive effects of CU traits and program involvement on antisocial outcomes (Hyp. 3)*

	<i>Panel A: T1 Mothers</i>					<i>Panel B: T1 Sons</i>				
Variable	R ²	β	SE	<i>t</i>	<i>p</i>	R ²	β	SE	<i>t</i>	<i>p</i>
	T1 Delinquency					T1 Delinquency				
Age		-.30	.20	-1.52	.13		--	--	--	--
Mentoring		.07	.28	.25	.81		-.11	.22	-.49	.63
Activities	.38**	.15	.23	.64	.52	.39*	.36	.26	1.37	.17
CU Traits		.60	.14	4.15	< .01		.50	.18	2.86	< .01
CU x Ment	.42**	.21	.17	1.26	.21	.42*	.25	.30	.82	.41
CU x Act	.38**	.04	.14	.28	.78	.51**	.38	.23	1.62	.11
	T1 Aggression					T1 Aggression				
Age		-.20	.20	-1.04	.30		--	--	--	--
Mentoring		.14	.22	.65	.52		.13	.31	.42	.67
Activities	.27*	.05	.25	.18	.86	.38*	.12	.30	.40	.69
CU Traits		.51	.15	3.39	< .01		.53	.20	2.64	< .01
CU x Ment	.27*	-.02	.16	-.12	.90	.47**	.35	.27	1.30	.20
CU x Act	.27*	-.08	.16	-.51	.61	.50**	.38	.24	1.58	.11

Note: * $p < .05$; ** $p < .01$

Table 10. *Hierarchical regression models estimating interaction effects of program involvement and callous-unemotional traits on antisocial outcomes at T2 (Hyp 3, T2)*

Variable	R ²	β	SE	t	p
T2 Delinquency					
Age		-.53	.29	-1.87	< .10
T2 Mentoring		-.01	.27	-.04	.97
T2 Activities	.32^t	-.29	.20	-1.41	.16
CU Traits		.05	.24	.19	.85
CU x T2 Ment	.32^t	-.04	.38	-.11	.92
CU x T2 Act	.31[*]	.24	.32	.74	.52
T1 Delinquency	.88^{**}	1.08	.32	3.37	< .01
T2 Aggression					
Age		-.39	.24	-1.67	< .10
T2 Mentoring		.14	.20	.68	.50
T2 Activities	.28^t	-.23	.21	-1.11	.27
CU Traits		.20	.20	.99	.32
CU x T2 Ment	.29^t	.05	.38	.14	.89
CU x T2 Act	.25^t	.17	.30	.59	.56
T1 Aggression	.42[*]	.44	.38	1.14	.25

Note: ^t $p < .10$; * $p < .05$; ** $p < .01$

Hypothesis 4. The fourth and final hypothesis stated that exposure to dangerous or stressful events would moderate associations between program participation and antisocial/academic outcomes, such that higher levels of program involvement would be associated with fewer antisocial behaviors and better academic performance among boys exposed to more stressors. In contrast, boys exposed to fewer stressors were expected to show weaker associations between program involvement and antisocial/academic outcomes.

Regression results at T1 are presented in Table 11 and T2 results are presented in Table 12.

Time 1. Bivariate associations at T1 revealed a marginally positive association between Stressors/Trauma and Physical Aggression ($r = .23, p = .09$), Reactive Aggression ($r = .26, p = .07$), and Proactive Aggression ($r = .23, p = .09$), as well as a significantly positive association

with the Total Aggression composite ($r = .30, p < .05$) among mother-reported data. Among youth-report, Stressors/Trauma were marginally positively related to Physical Aggression ($r = .29; p = .08$). These associations indicate a slight trend towards higher levels of aggressive behavior among boys exposed to more stressful and/or traumatic events and is consistent with existing literature on the subject (e.g., see Widom, 1989). As shown in Table 11, Panel A, Trauma/Stressors were significantly and positively associated with Aggression beyond the effects of either Mentoring or Activity Intensity ($\beta = .31, p < .05$), indicating that boys exposed to more traumatic or stressful events exhibited more mother-reported aggressive behavior. No significant effects of Mentoring, Activity Intensity, or Trauma/Stressors were found for any other mother- or youth-report models. No significant Trauma/Stressors x Activity Intensity/Mentoring interactions were found for mother- or youth report of Delinquency or Aggression at T1, and no main or interactive effects of Trauma/Stressors emerged in mother- or youth-report models predicting GPA.

Time 2. Bivariate analyses at T2 did not reveal any significant associations between Stressors/Trauma and main study variables (see Table 5). Regression results at T2 are presented in Table 12. Cross-sectional analyses of the effect of Trauma/Stressors on academic and antisocial outcomes at T2 revealed a significant and negative main effect of Activity Breadth on T2 Delinquency ($\beta = -.34, p < .05$) beyond the effect of Trauma/Stressors. This finding indicates that, in the current sample, greater breadth of activity participation was associated with less delinquency at T2 regardless of boys' exposure to trauma. No other significant main effects were detected for Aggression or GPA. No significant interactions between Trauma/Stressors x T2 Activities or Trauma/Stressors x T2 Mentoring emerged on models estimating T2 Delinquency, T2 Aggression, or T2 GPA.

Table 11. *Regression analyses estimating main and interactive effects of exposure to stressors/trauma on antisocial and academic outcomes (Hyp. 4)*

	<i>Panel A: T1 Mothers</i>					<i>Panel B: T1 Sons</i>				
Variable	R²	β	SE	t	p	R²	β	SE	t	p
	T1 Delinquency					T1 Delinquency				
Age		-.35	.21	-1.63	.10		--	--	--	--
Mentoring	.16	.03	.29	.09	.93	.18	-.02	.29	-.08	.93
Activities		-.17	.25	-.70	.48		.36	.34	1.07	.29
Trauma		.19	.18	1.06	.29		.17	.21	.80	.42
Trau x Ment		.16	.05	.21	.25		.80	.22	.26	.31
Trau x Act	.16	.01	.24	.03	.98	.18	-.12	.29	-.40	.69
	T1 Aggression					T1 Aggression				
Age		-.24	.19	-1.26	.21		--	--	--	--
Mentoring	.18	.09	.25	.36	.72	.14	.22	.41	.53	.59
Activities		-.24	.24	-1.00	.32		.13	.42	.31	.76
Trauma		.31	.15	2.05	< .05		.14	.21	.68	.50
Trau x Ment		.18	.02	.20	.07		.94	.20	.31	.31
Trau x Act	.18	-.02	.22	-.07	.94	.14	-.01	.29	-.02	.99
	T1 GPA					T1 GPA				
Age		-.25	.17	-1.43	.15		--	--	--	--
Mentoring	.11	-.11	.32	-.34	.73	.07	.20	.29	.71	.48
Activities		.20	.29	.71	.48		-.11	.27	-.39	.70
Trauma		-.20	.15	-1.37	.17		-.22	.24	-.89	.37
Trau x Ment		.11	.03	.17	.16		.88	.09	-.30	.26
Trau x Act	.12	.04	.17	.24	.81	.16	-.41	.33	-1.25	.21

Table 12. *Hierarchical regression analyses estimating interaction effects of program involvement and trauma/stressors on antisocial and academic outcomes at T2 (Hyp 4, T2).*

Variable	R ²	β	SE	t	p
T2 Delinquency					
Age		-.51	.27	-1.91	< .10
T2 Mentoring		-.03	.26	-.11	.91
T2 Activities	.34^t	-.34	.19	-1.79	< .10
Trauma		.17	.21	.84	.40
Trauma x T2 Ment	.35^t	.12	.38	.81	.42
Trauma x T2 Act	.35[*]	.09	.28	.32	.75
T1 Delinquency	.78^{**}	.80	.16	4.96	< .01
T2 Aggression					
Age		-.43	.23	-1.87	< .10
T2 Mentoring		.14	.21	.67	.50
T2 Activities	.26^t	-.29	.21	-1.39	.17
Trauma		-.10	.22	-.44	.66
Trauma x T2 Ment	.26^t	-.06	.36	-.16	.88
Trauma x T2 Act	.29[*]	.21	.24	.88	.38
T1 Aggression	.44^{**}	.49	.31	1.62	.11
T2 GPA					
Age		-.02	.28	-.06	.96
T2 Mentoring		-.40	.21	-1.86	< .10
T2 Activities	.19	.23	.24	.96	.34
Trauma		-.15	.28	-.52	.61
Trauma x T2 Ment	.21	.18	.44	.40	.69
Trauma x T2 Act	.21	.15	.34	.43	.67
T1 GPA	.51^{**}	.75	.29	2.64	< .01

Note: ^t $p < .10$; * $p < .05$; ** $p < .01$

Supplementary analyses. Given that the small sample size prevented the use of more complex longitudinal analyses, a series of exploratory moderations were tested as an alternative means of assessing change over time. Specifically, Activities and Mentoring were tested as moderators of associations between Delinquency, Aggression, and GPA to determine whether program participation affected boys' trajectories on antisocial and academic outcomes. Results indicated no significant interactions between Activities or Mentoring and T1-T2 antisocial or academic

outcomes. However, this finding is unsurprising, given the lack of significant differences detected on outcome variables between the two assessment points.

Discussion

The Son of a Saint youth organization is a mentorship-based program, which seeks to provide positive male role models for fatherless young boys. The program seeks to teach boys the importance of academic performance, as well as developing and maintaining strong ties to the community, and encourages positive relationships with peers through participation in recreational, vocational, and life skills training activities. The current pilot study sought to obtain preliminary information regarding the effectiveness of different aspects of SOAS programming at improving boys' behavior and academic performance. The primary aims of the study were to a) describe the intensity and breadth of boys' activity participation and participation in 1:1 mentoring, b) to examine the effects of program involvement on boys' antisocial (i.e., delinquent/aggressive) behavior and academic performance, both cross-sectionally and longitudinally, and c) examine the potential moderating role of individual (i.e., CU traits) and environmental (i.e., trauma/stress exposure) factors in associations between program involvement and antisocial/academic outcomes.

Although many evaluation studies are available in the current literature, they often do not include comprehensive assessment of contextual factors among program participants that may influence the effectiveness of program components such as mentorship and group activity participation (Roth & Brooks-Gunn, 2016). The current study is one of very few examining the influence of children's environment (i.e., exposure to trauma/stressors) on levels of program involvement and associated outcomes (see Mahoney, Vandell, Simpkins, & Zarrett, 2009, Roth & Brooks-Gunn, 2016, for reviews). In addition, this study is the first known program evaluation to include CU traits as a potential moderator of program outcomes.

Several patterns emerged from this data that warrant continued, comprehensive evaluation of the SOAS program in order to obtain a clearer picture of program effects. The following sections will be presented in two parts. The first will discuss youth participation in mentoring and activities at SOAS, how each relate to the antisocial and academic outcomes hypothesized, and the degree to which associations between program involvement and behavioral outcomes are affected by characteristics of the child or his environment. The second will address study limitations and recommendations for future directions that can help to bridge the divide between developmental and public health approaches to evaluation.

Aim 1: Program Involvement

Dosage, or the amount of time spent participating in various aspects of a program, is a vital component of program effectiveness (e.g., Larson & Verma, 1999; Simpkins, Little, & Weiss, 2004). In the same way that higher doses of a drug/medication produce more pronounced effects than low doses, higher levels of participation in program activities will ostensibly lead to stronger or more detectable effects on youth behavior and overall outcomes (e.g., Lochman, Boxmeyer, Powell, Roth, & Windle, 2006). Put differently, in order for youth to obtain the maximum benefits of a given program, they must participate on a regular basis and be engaged with program components (e.g., program activities; Ehrlich et al., 2017). In the current study, dosage was measured by duration of program involvement, the intensity and breadth of group activity participation, and 1:1 mentoring sessions. Given the wide range of duration of involvement across the full sample (3-64 months), duration was examined both as a predictor of antisocial/academic outcomes and as a potential moderating factor in associations between study outcome variables and participation in activities/mentoring.

Mentoring. Mothers at T1 reported that boys spent an average of 2.8 hours per week in 1:1 mentoring, while boys reported an average of 3.5 hours with mentors. At the 1-year follow-up assessment (T2), mothers reported boys were spending an average of 5 hours with mentors each week, representing a statistically significant increase in mentoring hours over the past year. Communication between the researchers and SOAS staff members revealed that the program had increased the number of mentors from 3-4 during the previous summer to 10-12 by the time of the T2 assessment, leading to greater availability of mentors to spend in 1:1 time with mentees. This increase in the mentor-mentee ratio may have important implications for boys' success as they continue in the SOAS program. Specifically, numerous studies have found that higher staff-to-youth ratios are associated with higher levels of youth-reported staff attention and support, less conflict among youth, and higher levels of satisfaction with the overall program experience (Hansen & Larson, 2007; Miller-Whitehead, 2003; Warren, Feist, & Nevarez, 2002). Thus, although the present study did not conduct additional assessments following the T2 survey, it will be important to examine whether the increase in number of mentors will be related to changes in boys' behavior, academic performance, and overall program involvement over time.

Activity participation. As discussed previously, activity participation at T1 was assessed according to intensity, or time spent participating in activities, while T2 assessed the breadth of different activities in which boys participated in over the past year. At T1, mothers reported that boys were spending an average of 3.5 hours per week participating in program activities, while boys reported an average of 4 hours per week. Data obtained from SOAS staff upon completion of T2 survey collection showed that boys had participated in an average of 13 different activities over the past year, with a range of 2-27 past-year activities across the full sample. Son of a Saint asks that boys participate in at least 2 activities each month but does not explicitly require them

to do so. When contacted by the research team for clarification, SOAS staff confirmed that some boys attend only the required annual fundraising gala and occasional Etiquette Meals. This finding is somewhat inconsistent with families' self-reported number of average weekly activity hours, though there are several possible explanations for the discrepancy. For example, it is possible that families were unable to accurately recall the average number of weekly hours boys spent participating. Alternatively, families may have provided the average hours spent participating in program activities for weeks in which they attended any SOAS-related events, rather than average weekly activity hours throughout the past year. In either case, it is clear that a more thorough and standardized metric (e.g., official records compiled by SOAS staff) for assessing boys' level of activity participation is needed in order to provide additional support for the findings in the current study.

Duration. Correlational data among youth self-report at T1 indicated that boys who have been in the program for shorter durations (i.e., less than 3 years) spend significantly more time in 1:1 mentoring sessions, and somewhat more time participating in other program activities. Similarly, mother-report at T1 showed that shorter duration of involvement was related to somewhat higher intensity of activity participation compared to longer duration of involvement. This pattern suggests that boys are more active in the program when they are new members and tend to participate less once they have been in the program for a longer period of time. Although the available data prevents statistical examination of this relationship, it is possible that boys participate less over time due to decreased need or increased time constraints once boys reach high school. Denault and Poulin (2009) provide support for this argument in their 5-year longitudinal assessment of rates of change in organized youth activity participation from grades 7 through 11. Specifically, the authors reported that youth participate in organized activities (e.g.,

sports, school-based, and community-oriented clubs or organizations) less frequently over time, with steeper declines in later adolescence. Such declines are expected as adolescents' undergo changes in autonomy, developmental priorities (e.g., romantic relationships), and time allocation (Fredricks & Eccles, 2006; Mahoney, Harris, & Eccles, 2006; Pedersen, 2005). Thus, declining rates of participation in activities and mentoring for boys who have been part of SOAS for longer periods of time is likely representative of normal developmental processes, rather than a decrease in the effectiveness of SOAS programming among older boys. However, the reduction in participation may be due to a number of other factors, including dissatisfaction with activities or incompatible mentor-mentee matches (e.g., Roth & Brooks-Gunn, 2016). Thus, a more comprehensive assessment of boys' motivations for participation at SOAS during adolescence is needed in order to corroborate findings from prior research.

Summary: Aim 1. Taken together, there is substantial variability across the full sample in both intensity and breadth of activity participation, as well as time spent in 1:1 mentoring sessions. This pattern is largely consistent with other youth NPOs, which often allow children and families to participate as much or as little as they choose, rather than enforcing strict guidelines for treatment adherence (Love et al., 2016; Roth & Brooks-Gunn, 2017). While this provides substantially more flexibility to accommodate families' needs and schedules, it may influence the degree to which some participants benefit from the program, depending on their level of need. Though the current data does not include an assessment of individual family needs, the following section will further explore the relationship between program involvement and key study outcomes.

Aim 2: Associations between program involvement and antisocial/academic outcomes

Mentoring and associated outcomes. At the bivariate level, participation in 1:1 mentoring sessions was positively related to T1 delinquency and aggression, and negatively related to T2 GPA. Time 2 Mentoring was significantly related to lower T2 GPA, even when accounting for the effects of age, activity participation, and duration of involvement at SOAS. In addition, longer duration of involvement was related to significantly higher GPA in bivariate analyses. This pattern of findings likely reflects increased need for 1:1 mentoring time for boys who continue to struggle academically even after they have been involved with SOAS for over a year. These results appear to support the previously suggested notion that boys with greater need for resources provided by SOAS are using these resources more than boys with lower need. The finding that mentoring was related to lower GPA regardless of boys' duration of involvement with the program is not surprising, given the emphasis SOAS places on improving academic achievement for boys in the program. That is, SOAS has strict academic requirements that boys must meet in order to maintain eligibility to participate (e.g., minimum GPA), thus most boys who have been in the program longer will have likely resolved any academic issues present when they were new to the program. It is also worth noting that these academic requirements may have resulted in a restricted range of academic outcomes in the current sample. Thus, additional descriptive data on mentor-mentee matching and continued follow-up assessments are needed to fully explore the relationship between boys' academic performance and time spent in 1:1 mentoring sessions. Examination of academic performance prior to boys' entry into SOAS will also help to clarify these associations.

Neither set of associations (i.e., between delinquency/aggression or academics and 1:1 mentoring) was consistent across time points, which is likely due to the change in number of

mentors at SOAS between T1 and T2 assessments. As noted previously, higher staff to youth ratios typically mean that mentors can devote more time to mentees who need more personalized or intensive assistance with behavioral or academic problems. Further, a higher number of available mentors increases the odds of mentees being matched with a mentor who is well-suited to the individual needs of the mentee. For example, a mentor with prior teaching or tutoring experience can be matched with boys who need more help academically (see Hansen & Larson, 2007; Tolan et al., 2013). Although new mentors receive general training regarding policies and procedures at SOAS, they receive less formal training in academic assistance (e.g., tutoring) or management of child behavior problems. It is therefore not always possible to match mentees with mentors who have adequate experience to support the individual needs of each boy in the program. Within the context of the current study, the higher levels of antisocial behavior associated with mentoring at T1 may indicate that mentees were not able to be matched with mentors who were equipped to adequately address boys' behavioral issues. On the other hand, the negative association between mentoring and GPA at T2 may indicate that boys who are performing poorly in school are being matched and spending more time with mentors who are suited to address boys' individual academic needs. Further, given that the average weekly mentoring hours significantly increased between T1 and T2 while delinquency, aggression, and GPA showed no significant changes, it is reasonable to suggest that the inconsistent associations with mentoring and antisocial/academic outcomes across assessment points may have been due to changes in mentor availability and/or the quality of mentor-mentee matches. It is unlikely that changes in behavioral or academic outcomes would become evident immediately when boys are matched with mentors who are qualified to address mentees' individual needs.

Activity participation and associated outcomes. Bivariate associations of mother-report at T1 indicated that both delinquency and aggression were positively related to the intensity of boys' activity participation, while T2 results revealed a trend toward lower levels of delinquency and aggression associated with the breadth of past-year activity participation. When duration of families' involvement at SOAS was included in the model as a covariate, mother-reported data indicated that boys' intensity of activity participation was related to lower levels of aggression for boys who have been in the program for less than a year. While the current data prohibits a more detailed examination of this finding, future evaluation studies of SOAS should consider exploring whether the association between intensity of activity participation and aggression among new boys in the program varies based on type of aggression (i.e., reactive and proactive aggression). As noted previously, reactive aggression is often related to impairments in social functioning (see Card & Little, 2006), and may therefore influence boys' likelihood of participating in group activities. Thus, a more in-depth examination of the potential relationship between aggression, social functioning, and peer relationships may help future researchers disentangle the association between activity participation and aggressive behavior among boys who are new to SOAS.

Activity participation was not related to GPA in bivariate or regression analyses, which may suggest that intensity and breadth of activities is less likely than mentoring to produce detectable effects on academic performance. Without having data on the breadth of activities at T1 or the intensity of activity participation at T2, no inferences can be made regarding the relative importance of intensity or breadth in predicting antisocial or academic outcomes. However, Denault and Poulin (2009) found that higher initial levels of both intensity *and* breadth of participation in organized activities were associated with more positive developmental

outcomes in later adolescence. Specifically, participation in a wider variety of activities (i.e., breadth) was related to fewer risk-taking behaviors and higher levels of community engagement (Denault & Poulin, 2009). Although causal inferences cannot be made from the available data, it is possible that participating in a variety of different program activities gives boys a way to fill free time and may provide exposure to prosocial peers who discourage boys from engaging in antisocial behavior (e.g., Fredricks & Eccles, 2006; Mahoney & Stattin, 2000). Neither T2 model remained significant after controlling for levels of T1 Delinquency and T1 Aggression, indicating that prior levels of antisocial behavior are the strongest indicators of current antisocial behavior. This finding is consistent with a large body of literature regarding the prediction of delinquency and adult criminality (e.g., Loeber & Farrington, 2012; Moffitt, 2017).

Summary: Aim 2. Although the T1 findings were largely in contrast to the hypothesized direction of program effects on antisocial behavior, it is possible that boys with more severe behavior problems utilize SOAS resources more frequently in general than do boys with fewer behavior problems. Ample evidence supports the notion that youth with disruptive behavior problems use mental health services, including community-based NPOs, at a higher rate than youth with internalizing psychopathology (i.e., depression and anxiety) or youth without significant behavioral or emotional problems (e.g., Burns et al., 1995; Wu et al., 1999). Although it is not possible to determine the particular reasons for this finding in the current sample, a number of possible explanations may exist. For example, one reason that parents are likely to seek out services for children with disruptive behavior problems is due to higher levels of parental stress caused by children who exhibit disruptive or aggressive behavior at home or school, particularly among single-parent families (Williford, Calkins, & Keane, 2007). It is therefore possible that mothers in the current sample initially encourage boys who are more

disruptive at home to participate more in SOAS activities because they feel that boys will derive maximum benefits from more intensive participation, thus reducing overall parental stress.

Taken together, the data indicate that antisocial behaviors may be more strongly related to participation in activities, while academic performance may be more strongly related to mentoring, although results of the current study do not show conclusive evidence that mentoring or activities directly affect levels of antisocial behavior or academic performance. In other words, boys who were more aggressive or delinquent tended to report greater intensity and breadth of activity participation, while boys with low GPAs tended to participate in more 1:1 mentoring. The existing literature discussed above provides support for this pattern, in that youth with disruptive behavior problems often show behavioral improvements associated with greater engagement in program activities and are more likely to utilize community resources aimed at preventing or reducing antisocial behaviors. However, the finding that aggression was related to less intensive activity participation for boys who have been with SOAS for shorter durations warrants additional follow up to determine whether social functioning or type of aggression may contribute to this effect.

Similarly, extant literature has shown that youth who struggle academically are likely to derive more benefit from 1:1 mentoring than from other aspects of community-based programs or interventions, although the benefits of mentoring for improving academics are largely dependent on the length and quality of the mentor-mentee relationship (see Grossman & Rhodes, 2002, for a review). Since the current data does not include information on the quality of mentoring or length of time that mentees have been paired with their current mentors, it is not yet possible to disentangle the effects of mentoring on academic outcomes for boys in the present sample. Efforts are currently underway to develop a protocol for evaluating the mentor-mentee

matching process and quality of the mentor-mentee relationships among boys in the program, which will allow for a more complete assessment of the effectiveness of the mentorship model currently used by SOAS. Additional follow-up assessments of boys' behavioral and academic performance and collection of consistent, detailed activity data are needed to further elucidate the effects of activity participation on psychological, social, and academic functioning. It is critically important to examine the relationship between participation in group activities and aggressive or delinquent behavior in longitudinal studies given the potential for iatrogenic effects that may arise from associating with deviant peers (see Tolan et al., 2014, for a review).

Aim 3: The moderating role of individual and environmental factors

As noted by Roth and Brooks-Gunn (2017), individual characteristics of the child and his or her environment are likely to impact the effectiveness of any intervention, including those provided by community-based NPOs. However, contextual influences are rarely considered when evaluating the effectiveness of youth programs and are typically limited to peer, family, or school contexts (Roth & Brooks-Gunn, 2017). The present study sought to test the influence of two contextual factors that are known to influence treatment effectiveness. Callous-unemotional (CU) traits were selected based on a large body of research that has shown high levels of treatment resistance among CU youth. Although CU traits have recently received more attention in clinical research settings, there remains a general lack of awareness of the potential impact of CU traits on community-based NPO effectiveness. Exposure to trauma and stressors also was selected due to its known associations with regard to effective intervention and to highlight the importance of using trauma-informed methods when developing and implementing programs for youth in high risk environments. While no moderating effects were found in the current study, there is sufficient evidence to warrant continued consideration of these and other contextual

influences that may affect the ability of SOAS and other NPOs to provide effective resources for youth.

Callous-Unemotional traits. Correlational analyses of mother-reported data showed that CU traits were associated with lower intensity of activity participation at T1 and less breadth of activity participation at T2. Boys with high CU traits were also significantly less likely than low-CU boys to participate in 1:1 mentoring at T1. This pattern suggests that boys with high CU traits may simply be less interested in becoming involved with mentors or participating in program activities. Contrary to expectations, program involvement did not differentially affect boys with high versus low CU traits in regression models. CU traits predicted delinquency and aggression beyond the effects of either mentoring or activity participation in T1 models, although this pattern did not emerge at T2. As discussed previously, the lack of consistency across time points may be related to measurement differences in program involvement constructs between T1 and T2. These results are consistent with existing literature showing strong associations between CU traits and antisocial behavior among youth (see Frick et al., 2014).

It is unclear whether CU traits would moderate the association between program involvement and delinquent/aggressive outcomes if significant effects of either mentoring or activities were present. In other words, the small sample size and variance in the current study made it challenging to detect significant effects of program involvement, so it is difficult to determine whether CU traits are associated with differential outcomes for boys in the current sample. The true nature of the relationship between CU traits, antisocial behavior, and programming at SOAS will likely become clear with a larger sample, consistent measurement of program components, and continued follow-up assessment. While boys' levels of CU traits are unlikely to change over time (i.e., because they are considered relatively stable over the course of

development; Frick & Morris, 2004; Frick & White, 2008), it is worth examining whether high quality mentor-mentee relationships or the “right” mix of activities could reduce delinquent or aggressive behaviors that are often associated with high CU traits. Thus, although the intervention effects (i.e., mentoring and activities) were not significant in the present study, these findings are consistent with the notion that CU traits may be associated with less effective treatment outcomes (e.g., see Hawes, Price, & Dadds, 2014, for a review) and should continue to be evaluated in studies of youth NPO effectiveness.

Exposure to trauma/stressors. Exposure to a greater number of stressors or traumatic events was related to significantly higher levels of mother-reported aggression and slightly higher mother- and youth-reported physical aggression at T1 in bivariate analyses. Further, regression analyses revealed that exposure to more stressors or traumatic events was associated with higher mother-reported T1 aggression when accounting for the effects of mentoring or activities. These findings are consistent with prior research (e.g., Marsee, 2008; Stimmel, Cruise, Ford, & Weiss, 2014) and suggest that boys in the current sample who have been exposed to more dangerous, stressful, or traumatic events may exhibit more aggressive behaviors than boys who have been exposed to fewer stressors. It should be noted that trauma histories were obtained via mother-report only, due to concerns from several mothers and SOAS staff members that some of the questions may be too upsetting for the boys in the study. Consequently, the data does not capture boys’ experiences of stressors or trauma, nor does it include a measure of post-traumatic stress symptoms that could clarify the impact of stress or trauma on boys’ overall emotional and behavioral functioning. Thus, while the measures used in the present study did not include detailed trauma histories for each boy, these findings support the well-documented association between exposure to trauma and/or violence during childhood and engaging in

aggressive/violent behavior during adolescence and adulthood (e.g., Widom, 1989, see also Miller & Marsee, 2019). Finally, GPA was unrelated to trauma/stressors across informants and assessment points, indicating that any current or past traumatic events boys have experienced are not significantly influencing academic performance overall. However, as noted previously, a larger sample size and a longer assessment period are needed to determine whether these findings remain consistent in subsequent evaluations. All boys in the program have experienced at least one traumatic event, as they must have lost their father to death or incarceration in order to be eligible to join SOAS. Thus, it is important to consider the use of a trauma-informed approach when developing and implementing each aspect of the program to fully address the needs of this vulnerable population.

Summary: Aim 3. Although the findings regarding the proposed moderators of program effectiveness were largely non-significant, CU traits and traumatic/stressful events do appear to exert some influence over both antisocial and academic outcomes and therefore warrant continued assessment when evaluating youth NPOs. In the case of CU traits, program effects may be more difficult to detect, given the strong association between CU and delinquent/aggressive behavior. That is, because CU traits are often strongly predictive of antisocial behavior, boys with high CU traits may be less amenable to the effects of participation in mentoring or program activities. CU traits did not predict antisocial behavior in T2 models, which may be notable given the increase in number of mentors between assessments and that activity participation at T2 was measured according to breadth, rather than intensity, of involvement. Further evaluation of the impact of CU traits on SOAS program effectiveness will likely provide clarity on whether increased mentorship or breadth of activity participation would produce significant effects on CU traits in larger samples over a longer assessment period.

Although CU traits are relatively rare in community-based samples, the current study showed a surprisingly normal distribution on the CU trait measure, indicating that the sample as a whole exhibits above-average levels of CU traits. Thus, it is worth exploring whether certain program components (e.g., range of available activities) should be tailored to meet the needs of the boys participating in the SOAS program. Finally, further exploration of the effects of trauma/stressors on boys' antisocial and academic outcomes is warranted to ensure that mentees are matched with mentors who are equipped to identify and address any behavioral or academic problems that are potentially related to boys' trauma histories.

Limitations and Future Directions

The current study was a pilot evaluation of a community-based NPO which did not have a consistent framework for evaluating program effectiveness at the time of the first assessment, similar to many NPOs. The preliminary nature of the study design, findings, and limited resources available to both the research team and SOAS staff resulted in a number of limitations. Nonetheless, each limitation can be used to inform future evaluation studies of SOAS and other youth NPOs more broadly.

The most substantial limitations related to the current study were the small sample size, limited assessment points, and a lack of access to necessary data, all of which made detecting significant effects challenging. In particular, the sample size prohibited more complex analyses with additional variables that may have allowed for a more robust interpretation of results. That is, although the data set contained information on numerous contextual variables (e.g., peer relationships, symptoms of depression/anxiety, pubertal status), it was not possible to include these variables in the current analyses with so few participants at each time point. Further, certain variables (i.e., school attendance and perceived school support) had to be eliminated from

planned analyses due to insufficient variability across the sample. In addition, small sample sizes often result in insufficient power to detect significant or meaningful effects which can be generalized to a larger population (e.g., Faber & Fonseca, 2014). It is therefore difficult to determine whether the results presented in the current study are representative of all boys at SOAS or youth involved in similar community NPOs.

The second major limitation was the lack of longitudinal assessment points at which program involvement and associated outcomes were collected. This limitation is particularly salient from a developmental perspective, as boys in the SOAS program begin during early adolescence and remain involved until they graduate from high school. Adolescence is a developmental period marked by substantial and frequent change, meaning that program components that are deemed “effective” when boys are 13 or 14 years old may not show the same degree of effectiveness once boys are 16 or 17 years old. Furthermore, the nature of NPOs often includes frequent changes to programming and structure based on available resources, community needs, and program growth (Mahoney et al., 2009). SOAS was undergoing changes in how activity participation was tracked and increasing the number of mentors over the course of the study period. NPOs often evolve over time as they adapt to meet the needs of participants and manage resources (Mahoney et al., 2009). However, the nature of these changes created substantial challenges for determining whether the results of the present study can be reasonably expected to be replicated in future evaluation studies. Completion of a third assessment one year after the second assessment was completed may help researchers determine whether T2 results (collected after program changes had been implemented) remain stable or show evidence of change when compared to a T3 assessment. As such, youth NPOs should be evaluated while

accounting for program changes to assess whether program components are developmentally appropriate as youth progress through adolescence and into adulthood.

Future studies should attempt to include both baseline and regular follow-up assessments for all boys participating in the SOAS program. Doing so will eliminate selection bias, meaning that results will be representative of the SOAS population as a whole. Moreover, collecting assessments from all members of SOAS will resolve the issues related to small sample size, as SOAS currently has nearly 100 active participants in the program. Regular (e.g., annual) follow-up assessments regarding boys' program involvement, behavioral issues, academic performance, and social-emotional adjustment would allow researchers and evaluation teams to more reliably determine the extent to which these variables change over the course of boys' involvement at SOAS. To assess the long-term impacts of the SOAS program, researchers should attempt to collect annual follow-up assessments after boys graduate from the program (i.e., after graduating from high school).

A third key limitation was that information regarding boys' past activity participation did not include the frequency with which boys participated in activities that were offered on multiple occasions. For example, horseback riding was offered on four separate occasions during the study period, but the research team was only able to collect information about whether boys had participated in horseback riding at any point over the past year. Consequently, researchers were unable to determine the full extent of boys' participation in group activities. SOAS offers a wide range of recreational, vocational, life skills, and community service activities, and it will be useful for future evaluations to include information on both the breadth and intensity of boys' participation in each type of activity. This limitation highlights the challenge of translating academic research into public policy, as SOAS staff cannot be expected to know the current state

of child development literature in the absence of formal training on the subject (Bach-Mortensen & Montgomery, 2018; Love et al., 2016). Prior research has shown that rates of both positive and negative developmental outcomes can vary depending upon the relative frequency of participation in one activity type over another (Hansen, Skorupski, & Arrington, 2010; Larson, Hansen, & Moneta, 2006). For example, Larson and colleagues (2006) reported that sports and fine arts activities were related to development of goal-setting behaviors, problem-solving skills, and time management, though sports-related activities also were associated with higher stress. Conversely, service-based activities were associated with the development of greater teamwork and more positive relationships with peers and adults. Thus, it will be valuable in the future to consider ways to collect data necessary to determine the differential impact of activity type and frequency of participation on boys' developmental outcomes. Additionally, future assessments should include items regarding whether boys were participating in extracurricular activities or other local youth NPOs. Such information is critical, as the effectiveness of a particular organization cannot be properly determined if boys are participating in multiple programs.

Similarly, researchers were unable to collect data related to boys' participation in 1:1 mentoring time. As a result, analyses were restricted to retrospective self-report of mentoring hours, which likely reduced the reliability of results (e.g., Kamphaus & Frick, 2005). Furthermore, there is substantial evidence suggesting that the quality of mentor-mentee relationships is of critical importance for achieving the maximum benefits of mentorship. For example, in a meta-analytic review conducted by Tolan and colleagues (2014), the authors identified four key components of successful mentoring programs: 1) shared interests or common experiences between mentors-mentees which help to facilitate a trusting relationship; 2) mentors

who are able to teach or provide information that help mentees navigate challenges related to social, academic, and family/peer relationships; 3) mentors who are able and willing to advocate for mentees across a variety of contexts; and 4) mentors who provide emotional support to promote a sense of self-worth and confidence in mentees. In addition, the length of the mentor-mentee relationship can have significant effects on boys' behavioral and academic outcomes. Specifically, Grossman and Rhodes (2002) reported that mentor-mentee matches lasting 12 months or longer were associated with significant improvements in academic competence, relationship quality, and reduction of delinquent behaviors. In contrast, when mentor-mentee relationships were terminated after less than 6 months, mentees showed significant declines in self-worth and perceived academic competence (Grossman & Rhodes, 2002). It would therefore be useful for research and evaluation teams to explore best practices for training mentors on handling complex issues, as well as to examine factors that may influence the longevity of mentor-mentee relationships. Such practices will maximize the chances that mentor-mentee matches are appropriate, stable, and are not harmful to the social or emotional functioning of mentees. Further, mentees may be asked to provide confidential feedback regarding the level of satisfaction with mentors on a regular basis. Feedback items may include (but are not limited to) whether the mentee feels he receives enough personalized attention from mentors, the extent to which the mentor-mentee relationship lives up to expectations, whether mentees feel they can relate to mentors and receive emotional support, and the ways in which they feel mentors have been most helpful. Finally, the current study did not include items related to boys' change in levels of participation over time. Including items assessing specific reasons or situations that are likely to motivate boys to participate in activities or mentoring over an extended period of time will be useful to determine whether the decline in participation over time is due to changes in

need for services, shifting priorities, or dissatisfaction with program components. Assessment of underlying motivations for participation in mentoring and activities is particularly important for programs such as SOAS, which largely allows the families to determine boys' level of program involvement. While many of the above directions for future research may not be feasible due to limited resources, these recommendations may help to provide a framework for designing future research and evaluation studies, both for SOAS and youth NPOs in general.

Conclusion

Evaluation of community-based NPOs is vital for determining whether components of these programs are effectively addressing the needs of participants and the goals of the organization. Consistent and comprehensive evaluation of youth NPOs is particularly important to ensure that program components are developmentally appropriate and consider the contextual aspects of children's lives that may influence treatment outcomes (e.g., Lochman et al., 2017; Roth & Brooks-Gunn, 2016). The current study provides preliminary data on how program participation is affecting behavioral and academic outcomes among participants while also highlighting the challenges of evaluation in community settings. Results of the study hypotheses were somewhat inconclusive but nonetheless inform ways to improve data collection and ongoing tracking of youth outcomes to more accurately assess the effectiveness of program objectives. Furthermore, approaching evaluation from both a developmental perspective and from a public health perspective can begin to close the gap between research and practice in real-world settings and will allow future evaluators and researchers to conduct contextually relevant and developmentally sensitive evaluations on the effectiveness of SOAS and related programs.

References

- Ammerman, R. T., Cassisi, J. E., Hersen, M., & Van Hasselt, V. B. (1986). Consequences of physical abuse and neglect in children. *Clinical psychology review*, 6(4), 291-310.
- Anda, R. F., Felitti, V. J., Bremner, J. D., Walker, J. D., Whitfield, C., Perry, B. D., . . . Giles, W. H. (2006). The enduring effects of abuse and related adverse experiences in childhood. *European archives of psychiatry and clinical neuroscience*, 256(3), 174-186.
- Amed, M. M. (2013). The efficacy of the Girls on the Run program to improve self-worth, body image, and behavioral and emotional functioning: A longitudinal study.
- Anderson, E. (1999). *Code of the Street*: New York: Norton.
- Anderson-Butcher, D., Newsome, W. S., & Ferrari, T. M. (2003). Participation in boys and girls clubs and relationships to youth outcomes. *Journal of Community Psychology*, 31(1), 39-55.
- Aos, S., Phipps, P., Barnoski, R., & Lieb, R. (2001). The Comparative Costs and Benefits of Programs to Reduce Crime. Version 4.0.
- Arnold, M. E., & Hughes, J. N. (1999). First do no harm: Adverse effects of grouping deviant youth for skills training. *Journal of School Psychology*, 37(1), 99-115.
- American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders*.
- Bach-Mortensen, A. M., & Montgomery, P. (2018). What are the barriers and facilitators for third sector organisations (non-profits) to evaluate their services? A systematic review. *Systematic reviews*, 7(1), 13.
- Barry, M. M., Clarke, A. M., Morreale, S. E., & Field, C. A. (2018). A review of the evidence on the effects of community-based programs on young people's social and emotional skills development. *Adolescent Research Review*, 3(1), 13-27.
- Belsky, J., & Pluess, M. (2009). Beyond diathesis stress: Differential susceptibility to environmental influences. *Psychological bulletin*, 135(6), 885.
- Berman, S. L., Kurtines, W. M., Silverman, W. K., & Serafini, L. T. (1996). The impact of exposure to crime and violence on urban youth. *American journal of Orthopsychiatry*, 66(3), 329-336.
- Bandura, A. (1972). Modeling theory: Some traditions, trends, and disputes. *Recent trends in social learning theory*, 35-61.
- Becker, H. K., Agopian, M. W., & Yeh, S. (1992). Impact evaluation of drug abuse resistance education (DARE). *Journal of drug education*, 22(4), 283-291.
- Bernat, D. H., Oakes, J. M., Pettingell, S. L., & Resnick, M. (2012). Risk and direct protective factors for youth violence: results from the National Longitudinal Study of Adolescent Health. *American journal of preventive medicine*, 43(2), S57-S66.
- Bialeschki, M. D., & Conn, M. (2011). Welcome to our world: Bridging youth development research in nonprofit and academic communities. *Journal of Research on Adolescence*, 21(1), 300-306.
- Birch, S. H., & Ladd, G. W. (1997). The teacher-child relationship and children's early school adjustment. *Journal of School Psychology*, 35(1), 61-79.
- Blakeslee, J. E., & Keller, T. E. (2018). *Extending a randomized trial of my life mentoring model for youth in foster care to evaluate long-term effects on offending in young adulthood: Final technical report*: Portland State University.
- Boat, B., Baker, D., & Abrahamson, S. (1996). The childhood trust events survey. *Department of Psychiatry*.

- Bolton, K. W., Maume, M. O., Jones Halls, J., & Smith, S. D. (2017). Multisectoral approaches to addressing youth violence. *Journal of Human Behavior in the Social Environment*, 27(7), 760-767.
- Brendgen, M., Vitaro, F., Tremblay, R. E., & Lavoie, F. (2001). Reactive and proactive aggression: Predictions to physical violence in different contexts and moderating effects of parental monitoring and caregiving behavior. *Journal of abnormal child psychology*, 29(4), 293-304.
- Brier, N. (1995). Predicting antisocial behavior in youngsters displaying poor academic achievement: A review of risk factors. *Journal of Developmental & Behavioral Pediatrics*, 16(4), 271-276.
- Broidy, L. M., Nagin, D. S., Tremblay, R. E., Bates, J. E., Brame, B., Dodge, K. A., . . . Laird, R. (2003). Developmental trajectories of childhood disruptive behaviors and adolescent delinquency: a six-site, cross-national study. *Developmental psychology*, 39(2), 222.
- Burns, B. J., Costello, E. J., Angold, A., Tweed, D., Stangl, D., Farmer, E. M., & Erkanli, A. (1995). Children's mental health service use across service sectors. *Health affairs*, 14(3), 147-159.
- Burt, S. A., & Donnellan, M. B. (2009). Development and validation of the Subtypes of Antisocial Behavior Questionnaire. *Aggressive Behavior: Official Journal of the International Society for Research on Aggression*, 35(5), 376-398.
- Busseri, M. A., Rose-Krasnor, L., Willoughby, T., & Chalmers, H. (2006). A longitudinal examination of breadth and intensity of youth activity involvement and successful development. *Developmental psychology*, 42(6), 1313.
- Busseri, M. A., & Rose-Krasnor, L. (2009). Breadth and intensity: Salient, separable, and developmentally significant dimensions of structured youth activity involvement. *British journal of developmental psychology*, 27(4), 907-933.
- Callahan, K. L., Scaramella, L. V., Laird, R. D., & Sohr-Preston, S. L. (2011). Neighborhood disadvantage as a moderator of the association between harsh parenting and toddler-aged children's internalizing and externalizing problems. *Journal of Family Psychology*, 25(1), 68.
- Cantor, P., Osher, D., Berg, J., Steyer, L., & Rose, T. (2018). Malleability, plasticity, and individuality: How children learn and develop in context¹. *Applied Developmental Science*, 1-31.
- Card, N. A., & Little, T. D. (2006). Proactive and reactive aggression in childhood and adolescence: A meta-analysis of differential relations with psychosocial adjustment. *International Journal of Behavioral Development*, 30(5), 466-480.
- Carman, J. G., & Fredericks, K. A. (2008). Nonprofits and evaluation: Empirical evidence from the field. *New Directions for Evaluation*, 2008(119), 51-71.
- Cavanagh, S. E., & Huston, A. C. (2008). The timing of family instability and children's social development. *Journal of Marriage and Family*, 70(5), 1258-1270.
- Cécile, M., & Born, M. (2009). Intervention in juvenile delinquency: Danger of iatrogenic effects? *Children and youth services review*, 31(12), 1217-1221.
- Chalmers, I. (2003). Trying to do more good than harm in policy and practice: the role of rigorous, transparent, up-to-date evaluations. *The Annals of the American Academy of Political and Social Science*, 589(1), 22-40.

- Coie, J. D., Watt, N. F., West, S. G., Hawkins, J. D., Asarnow, J. R., Markman, H. J., . . . Long, B. (1993). The science of prevention: A conceptual framework and some directions for a national research program. *American psychologist*, 48(10), 1013.
- Côté, J. (1999). The influence of the family in the development of talent in sport. *The sport psychologist*, 13(4), 395-417.
- Courser, M., Shamblen, S., Thompson, K., Young, L., Schweinhart, A., Shepherd, C., . . . Evaluation. (2017). *Assessing the impact of parental characteristics, parental attitudes, and parental engagement on mentoring relationship outcomes*: Pacific Institute for Research and Evaluation.
- Crapanzano, A. M., Frick, P. J., & Terranova, A. M. (2010). Patterns of physical and relational aggression in a school-based sample of boys and girls. *Journal of abnormal child psychology*, 38(4), 433-445.
- Crick, N. R., & Dodge, K. A. (1996). Social information-processing mechanisms in reactive and proactive aggression. *Child development*, 67(3), 993-1002.
- Curtis, N. M., Ronan, K. R., & Borduin, C. M. (2004). Multisystemic treatment: a meta-analysis of outcome studies. *Journal of Family Psychology*, 18(3), 411.
- DARE. (2019). The History of DARE. Retrieved from <https://dare.org/history/>
- De Castro, B. O., Veerman, J. W., Koops, W., Bosch, J. D., & Monshouwer, H. J. (2002). Hostile attribution of intent and aggressive behavior: A meta-analysis. *Child development*, 73(3), 916-934.
- Denault, A.-S., & Poulin, F. (2009). Intensity and breadth of participation in organized activities during the adolescent years: Multiple associations with youth outcomes. *Journal of Youth and Adolescence*, 38(9), 1199-1213.
- Despard, M. R. (2016). Challenges in implementing evidence-based practices and programs in nonprofit human service organizations. *Journal of evidence-informed social work*, 13(6), 505-522.
- Dishion, T. J., McCord, J., & Poulin, F. (1999). When interventions harm: Peer groups and problem behavior. *American psychologist*, 54(9), 755.
- Dishion, T. J., Spracklen, K. M., Andrews, D. W., & Patterson, G. R. (1996). Deviancy training in male adolescent friendships. *Behavior Therapy*, 27(3), 373-390.
- Dodge, K. A. (1991). *The structure and function of reactive and proactive aggression*. Paper presented at the Earls court Symposium on Childhood Aggression, Jun, 1988, Toronto, ON, Canada.
- Dodge, K. A., & Coie, J. D. (1987). Social-information-processing factors in reactive and proactive aggression in children's peer groups. *Journal of personality and social psychology*, 53(6), 1146.
- Dodge, K. A., Coie, J. D., & Lynam, D. (2006). Aggression and antisocial behavior in youth. *Handbook of child psychology*.
- Dodge, K. A., Lochman, J. E., Harnish, J. D., Bates, J. E., & Pettit, G. S. (1997). Reactive and proactive aggression in school children and psychiatrically impaired chronically assaultive youth. *Journal of abnormal psychology*, 106(1), 37.
- Dopp, A. R., Borduin, C. M., Wagner, D. V., & Sawyer, A. M. (2014). The economic impact of multisystemic therapy through midlife: A cost-benefit analysis with serious juvenile offenders and their siblings. *Journal of consulting and clinical psychology*, 82(4), 694.

- Drake, E. K., Aos, S., & Miller, M. G. (2009). Evidence-based public policy options to reduce crime and criminal justice costs: Implications in Washington State. *Victims and Offenders, 4*(2), 170-196.
- DuBois, D. L., Portillo, N., Rhodes, J. E., Silverthorn, N., & Valentine, J. C. (2011). How effective are mentoring programs for youth? A systematic assessment of the evidence. *Psychological Science in the Public Interest, 12*(2), 57-91.
- Eisenberg, N., Guthrie, I. K., Fabes, R. A., Shepard, S., Losoya, S., Murphy, B., . . . Reiser, M. (2000). Prediction of elementary school children's externalizing problem behaviors from attentional and behavioral regulation and negative emotionality. *Child development, 71*(5), 1367-1382.
- Eisenberg, N., & Spinrad, T. L. (2004). Emotion-related regulation: Sharpening the definition. *Child development, 75*(2), 334-339.
- Ennett, S. T., Tobler, N. S., Ringwalt, C. L., & Flewelling, R. L. (1994). How effective is drug abuse resistance education? A meta-analysis of Project DARE outcome evaluations. *American journal of public health, 84*(9), 1394-1401.
- Faber, J., & Fonseca, L. M. (2014). How sample size influences research outcomes. *Dental press journal of orthodontics, 19*(4), 27-29.
- Farrell, P. T., Woods, K. A., Squires, G., Rooney, S., Lewis, S., & O'Connor, M. (2007). Educational Psychologists' contribution to the Every Child Matters agenda: the parents' view. *Educational Psychology in Practice, 23*(4).
- Farrington, D. P. (1989). Early predictors of adolescent aggression and adult violence. *Violence and victims, 4*(2), 79.
- Farrington, D. P. (2003). Developmental and life-course criminology: Key theoretical and empirical issues-the 2002 Sutherland Award address. *Criminology, 41*(2), 221-225.
- Felson, R. B., & Staff, J. (2006). Explaining the academic performance-delinquency relationship. *Criminology, 44*(2), 299-320.
- Fergus, S., & Zimmerman, M. A. (2005). Adolescent resilience: A framework for understanding healthy development in the face of risk. *Annu. Rev. Public Health, 26*, 399-419.
- Finckenauer, J. (1982). *Scared Straight and the Panacea Problem*: Prentice-Hall.
- Finckenauer, J. O., Gavin, P. W., Hovland, A., & Storvoll, E. (1999). *Scared Straight!: The Panacea Phenomenon Revisited*: Waveland Press Prospect Heights, IL.
- Fletcher, A. C., & Shaw, R. A. (2000). Sex differences in associations between parental behaviors and characteristics and adolescent social integration. *Social Development, 9*(2), 133-148.
- Foster, W. L., Kim, P., & Christiansen, B. (2009). Ten nonprofit funding models.
- Fredricks, J. A., & Eccles, J. S. (2006). Is extracurricular participation associated with beneficial outcomes? Concurrent and longitudinal relations. *Developmental psychology, 42*(4), 698.
- Fredricks, J. A., Naftzger, N., Smith, C., & Riley, A. (2017). Measuring youth participation, program quality, and social and emotional skills in after-school programs *After-school programs to promote positive youth development* (pp. 23-43): Springer.
- Frick, P. J. (2004). The inventory of callous-unemotional traits. *Unpublished rating scale*.
- Frick, P. J., Cornell, A. H., Barry, C. T., Bodin, S. D., & Dane, H. E. (2003). Callous-unemotional traits and conduct problems in the prediction of conduct problem severity, aggression, and self-report of delinquency. *Journal of abnormal child psychology, 31*(4), 457-470.

- Frick, P. J., Cornell, A. H., Bodin, S. D., Dane, H. E., Barry, C. T., & Loney, B. R. (2003). Callous-unemotional traits and developmental pathways to severe conduct problems. *Developmental psychology, 39*(2), 246.
- Frick, P. J., & Dickens, C. (2006). Current perspectives on conduct disorder. *Current psychiatry reports, 8*(1), 59-72.
- Frick, P. J., & Morris, A. S. (2004). Temperament and developmental pathways to conduct problems. *Journal of Clinical Child and Adolescent Psychology, 33*(1), 54-68.
- Frick, P. J., Ray, J. V., Thornton, L. C., & Kahn, R. E. (2014). Can callous-unemotional traits enhance the understanding, diagnosis, and treatment of serious conduct problems in children and adolescents? A comprehensive review. *Psychological Bulletin, 140*(1), 1.
- Frick, P. J., & White, S. F. (2008). Research review: The importance of callous-unemotional traits for developmental models of aggressive and antisocial behavior. *Journal of Child Psychology and Psychiatry, 49*(4), 359-375.
- Furlong, M. J., You, S., Renshaw, T. L., Smith, D. C., & O'Malley, M. D. (2014). Preliminary development and validation of the social and emotional health survey for secondary school students. *Social Indicators Research, 117*(3), 1011-1032.
- Gabriel, K. K. P., DeBate, R. D., High, R. R., & Racine, E. F. (2011). Girls on the Run: a quasi-experimental evaluation of a developmentally focused youth sport program. *Journal of Physical Activity and Health, 8*(s2), S285-S294.
- Gorman-Smith, D., Tolan, P. H., & Henry, D. (1999). The relation of community and family to risk among urban-poor adolescents. *Historical and geographical influences on psychopathology, 349-367*.
- Greenwood, P. W., & Turner, S. (2011). Establishing Effective Community-Based Care in Juvenile Justice. *Juvenile Justice: Advancing Research, Policy, and Practice, 477-504*.
- Grossman, J. B., & Garry, E. M. (1997). *Mentoring: A proven delinquency prevention strategy*: US Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention Washington, DC.
- Grossman, J. B., & Rhodes, J. E. (2002). The test of time: Predictors and effects of duration in youth mentoring relationships. *American journal of community psychology, 30*(2), 199-219.
- Grossman, J. B., & Tierney, J. P. (1998). Does mentoring work? An impact study of the Big Brothers Big Sisters program. *Evaluation review, 22*(3), 403-426.
- Guerra, N. G., & Huesmann, L. R. (2004). A Cognitive-Ecological Theory of Aggression. *Revue Internationale de Psychologie Sociale*.
- Hanlon, T. E., Simon, B. D., O'Grady, K. E., Carswell, S. B., & Callaman, J. M. (2009). The effectiveness of an after-school program targeting urban african american youth. *Education and Urban Society, 42*(1), 96-118.
- Hansen, D. M., & Larson, R. W. (2007). Amplifiers of developmental and negative experiences in organized activities: Dosage, motivation, lead roles, and adult-youth ratios. *Journal of applied developmental psychology, 28*(4), 360-374.
- Hansen, D. M., Skorupski, W. P., & Arrington, T. L. (2010). Differences in developmental experiences for commonly used categories of organized youth activities. *Journal of Applied Developmental Psychology, 31*(6), 413-421.
- Hart-Johns, M., Courser, M., Kirk, H., Research, P. I. f., Evaluation, & America, U. S. o. (2017). Lessons Learned and Implications from a Cross-Site Evaluation of Mentoring With System-Involved Youth.

- Hawes, D. J., & Dadds, M. R. (2005). The treatment of conduct problems in children with callous-unemotional traits. *Journal of consulting and clinical psychology, 73*(4), 737.
- Hawes, D. J., Price, M. J., & Dadds, M. R. (2014). Callous-unemotional traits and the treatment of conduct problems in childhood and adolescence: A comprehensive review. *Clinical child and family psychology review, 17*(3), 248-267.
- Henggeler, S., Schoenwald, S., Borduin, C., Rowland, M., & Cunningham, P. (1998). Treatment manuals for practitioners: Multisystemic treatment of antisocial behavior in children and adolescents
- Henggeler, S., Schoenwald, S., Borduin, C. M., Rowland, M., & Cunningham, P. (2009). Multisystemic therapy for antisocial behaviour in children and adolescents. *New York, NY: Guilford.*
- Henggeler, S. W., & Borduin, C. M. (1990). *Family therapy and beyond: A multisystemic approach to treating the behavior problems of children and adolescents*: Thomson Brooks/Cole.
- Henggeler, S. W., & Schaeffer, C. (2010). Treating serious antisocial behavior using multisystemic therapy. *Evidence-based psychotherapies for children and adolescents, 2*, 259-276.
- Henggeler, S. W., Schoenwald, S. K., Borduin, C. M., Rowland, M. D., & Cunningham, P. B. (2009). *Multisystemic therapy for antisocial behavior in children and adolescents*: Guilford Press.
- Herman, R. D., & Renz, D. O. (1998). Nonprofit organizational effectiveness: Contrasts between especially effective and less effective organizations. *Nonprofit management and leadership, 9*(1), 23-38.
- Hoeve, M., Dubas, J. S., Eichelsheim, V. I., Van Der Laan, P. H., Smeenk, W., & Gerris, J. R. (2009). The relationship between parenting and delinquency: A meta-analysis. *Journal of abnormal child psychology, 37*(6), 749-775.
- Holmes, S. E., Slaughter, J. R., & Kashani, J. (2001). Risk factors in childhood that lead to the development of conduct disorder and antisocial personality disorder. *Child Psychiatry and Human Development, 31*(3), 183-193.
- Holt, M. K., Finkelhor, D., & Kantor, G. K. (2007). Multiple victimization experiences of urban elementary school students: Associations with psychosocial functioning and academic performance. *Child abuse & neglect, 31*(5), 503-515.
- Jolliffe, D., Farrington, D. P., Loeber, R., & Pardini, D. (2016). Protective factors for violence: Results from the Pittsburgh Youth Study. *Journal of Criminal Justice, 45*, 32-40.
- Kamphaus, R. W., & Frick, P. J. (2005). *Clinical assessment of child and adolescent personality and behavior*: Springer Science & Business Media.
- Karcher, M. J., & Nakkula, M. J. (2010). Youth mentoring with a balanced focus, shared purpose, and collaborative interactions. *New directions for youth development, 2010*(126), 13-32.
- Keating, L. M., Tomishima, M. A., Foster, S., & Alessandri, M. (2002). The effects of a mentoring program on at-risk youth. *ADOLESCENCE-SAN DIEGO-*, *37*, 717-734.
- Kim, B. K. E., Gilman, A. B., & Hawkins, J. D. (2015). 28 School- and Community-Based Preventive Interventions During Adolescence: Preventing Delinquency Through Science-Guided Collective Action. In J. Morizot & L. Kazemian (Eds.), *The Development of Criminal and Antisocial Behavior: Theory, Research and Practical Applications* (pp. 447-460). Cham: Springer International Publishing.

- Kim, H. K., & Leve, L. D. (2011). Substance use and delinquency among middle school girls in foster care: A three-year follow-up of a randomized controlled trial. *Journal of consulting and clinical psychology, 79*(6), 740.
- Krueger, R. F., Markon, K. E., Patrick, C. J., Benning, S. D., & Kramer, M. D. (2007). Linking antisocial behavior, substance use, and personality: an integrative quantitative model of the adult externalizing spectrum. *Journal of abnormal psychology, 116*(4), 645.
- Larson, R. W., Hansen, D. M., & Moneta, G. (2006). Differing profiles of developmental experiences across types of organized youth activities. *Developmental psychology, 42*(5), 849.
- Larson, R. W., & Verma, S. (1999). How children and adolescents spend time across the world: Work, play, and developmental opportunities. *Psychological bulletin, 125*(6), 701.
- Larson, S., Chapman, S., Spetz, J., & Brindis, C. D. (2017). Chronic childhood trauma, mental health, academic achievement, and school-based health center mental health services. *Journal of school health, 87*(9), 675-686.
- Lerner, R. M., Dowling, E. M., & Anderson, P. M. (2003). Positive youth development: Thriving as the basis of personhood and civil society. *Applied Developmental Science, 7*(3), 172-180.
- Lewis, R. V. (1983). Scared straight—California style: Evaluation of the San Quentin SQUIRES program. *Criminal Justice and Behavior, 10*(2), 209-226.
- Limbos, M. A., Chan, L. S., Warf, C., Schneir, A., Iverson, E., Shekelle, P., & Kipke, M. D. (2007). Effectiveness of interventions to prevent youth violence: A systematic review. *American journal of preventive medicine, 33*(1), 65-74.
- Lipsey, M. W., Howell, J. C., Kelly, M. R., Chapman, G., & Carver, D. (2010). Improving the effectiveness of juvenile justice programs. *Washington DC: Center for Juvenile Justice Reform at Georgetown University.*
- Lochman, J. E., Boxmeyer, C., Powell, N., Roth, D. L., & Windle, M. (2006). Masked intervention effects: Analytic methods for addressing low dosage of intervention. *New Directions for Evaluation, 2006*(110), 19-32.
- Lochman, J. E., Dishion, T. J., Boxmeyer, C. L., Powell, N. P., & Qu, L. (2017). Variation in response to evidence-based group preventive intervention for disruptive behavior problems: a view from 938 coping power sessions. *Journal of abnormal child psychology, 45*(7), 1271-1284.
- Loeber, R. (1990). Development and risk factors of juvenile antisocial behavior and delinquency. *Clinical psychology review, 10*(1), 1-41.
- Loeber, R., & Farrington, D. P. (2000). Young children who commit crime: Epidemiology, developmental origins, risk factors, early interventions, and policy implications. *Development and psychopathology, 12*(4), 737-762.
- Loeber, R., & Farrington, D. P. (2012). *From juvenile delinquency to adult crime: Criminal careers, justice policy, and prevention*: Oxford university press.
- Loeber, R., & Stouthamer-Loeber, M. (1998). Development of juvenile aggression and violence: Some common misconceptions and controversies. *American psychologist, 53*(2), 242.
- Love, H., Harvell, S., & Winkler, M. K. (2016). *Understanding Research and Practice Gaps in Juvenile Justice.*
- Lundman, R. J. (2001). *Prevention and control of juvenile delinquency*: Oxford University Press.

- Lynam, D. R., Milich, R., Zimmerman, R., Novak, S. P., Logan, T., Martin, C., . . . Clayton, R. (1999). Project DARE: no effects at 10-year follow-up. *Journal of consulting and clinical psychology, 67*(4), 590.
- Maguin, E., & Loeber, R. (1996). Academic performance and delinquency. *Crime and justice, 20*, 145-264.
- Mahoney, J. L., Cairns, B. D., & Farmer, T. W. (2003). Promoting interpersonal competence and educational success through extracurricular activity participation. *Journal of educational psychology, 95*(2), 409.
- Mahoney, J. L., Harris, A. L., & Eccles, J. S. (2006). Organized activity participation, positive youth development, and the over-scheduling hypothesis. Social policy report. Volume 20, number 4. *Society for Research in Child Development*.
- Mahoney, J. L., & Stattin, H. (2000). Leisure activities and adolescent antisocial behavior: The role of structure and social context. *Journal of adolescence, 23*(2), 113-127.
- Mahoney, J. L., Vandell, D. L., Simpkins, S., & Zarrett, N. (2009). Adolescent out-of-school activities. *Handbook of adolescent psychology, 2*, 228-269.
- Majer, M., Nater, U. M., Lin, J.-M. S., Capuron, L., & Reeves, W. C. (2010). Association of childhood trauma with cognitive function in healthy adults: A pilot study. *BMC neurology, 10*(1), 61.
- Markham, A. (2018). A review following systematic principles of multisystemic therapy for antisocial behavior in adolescents aged 10–17 years. *Adolescent Research Review, 3*(1), 67-93.
- Marsee, M. A. (2008). Reactive aggression and posttraumatic stress in adolescents affected by Hurricane Katrina. *Journal of Clinical Child & Adolescent Psychology, 37*(3), 519-529.
- Marsee, M. A., Barry, C. T., Childs, K. K., Frick, P. J., Kimonis, E. R., Muñoz, L. C., . . . Lau, K. S. (2011). Assessing the forms and functions of aggression using self-report: Factor structure and invariance of the Peer Conflict Scale in youths. *Psychological assessment, 23*(3), 792.
- Marsee, M. A., & Frick, P. J. (2007). Exploring the cognitive and emotional correlates to proactive and reactive aggression in a sample of detained girls. *Journal of abnormal child psychology, 35*(6), 969-981.
- McLanahan, S., Tach, L., & Schneider, D. (2013). The causal effects of father absence. *Annual review of sociology, 39*, 399-427.
- Miller, M. A., & Marsee, M. A. (2019). Emotional Reactivity and Antisocial Behavior Relative to Posttraumatic Stress Symptom Expression: a Latent Profile Analysis. *Journal of abnormal child psychology*. [doi:10.1007/s10802-019-00514-9](https://doi.org/10.1007/s10802-019-00514-9)
- Miller-Whitehead, M. (November, 2003) *Compilation of class size findings: Grade level, school, and district*. Paper presented at the Annual Meeting of the mid-south educational research association, Biloxi, MS.
- Moffitt, T. E. (2017). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. In *Biosocial theories of crime* (pp. 69-96): Routledge.
- Murray, V., & Tassie, B. (1994). Evaluating the effectiveness of nonprofit organizations. *The Jossey-Bass Handbook of nonprofit leadership and management, 303-324*.
- Muthén, L. K., & Muthén, B. O. (2012). *Mplus: Statistical Analysis with Latent Variables; User's Guide;[version 7]*: Muthén et Muthén.

- Najaka, S. S., Gottfredson, D. C., & Wilson, D. B. (2001). A meta-analytic inquiry into the relationship between selected risk factors and problem behavior. *Prevention Science*, 2(4), 257-271.
- Norris-Tirrell, D. (2014). The changing role of private, nonprofit organizations in the development and delivery of human services in the United States. *Journal of health and human services administration*, 304-326.
- Office of Juvenile Justice and Delinquency Prevention. (2019). Model Programs Guide. Available from Office of Juvenile Justice and Delinquency Prevention OJJDP Model Programs Guide <https://www.ojjdp.gov/mpg>
- Overstreet, S., & Mathews, T. (2011). Challenges associated with exposure to chronic trauma: Using a public health framework to foster resilient outcomes among youth. *Psychology in the Schools*, 48(7), 738-754.
- Pardini, D. A., Loeber, R., Farrington, D. P., & Stouthamer-Loeber, M. (2012). Identifying direct protective factors for nonviolence. *American journal of preventive medicine*, 43(2), S28-S40.
- Pearl, E., Thieken, L., Olafson, E., Boat, B., Connelly, L., Barnes, J., & Putnam, F. (2012). Effectiveness of community dissemination of parent-child interaction therapy. *Psychological trauma: theory, research, practice, and policy*, 4(2), 204.
- Pedersen, S. (2005). Urban adolescents' out-of-school activity profiles: Associations with youth, family, and school transition characteristics. *Applied Developmental Science*, 9(2), 107-124.
- Petrosino, A., Turpin-Petrosino, C., & Buehler, J. (2003). Scared Straight and other juvenile awareness programs for preventing juvenile delinquency: A systematic review of the randomized experimental evidence. *The Annals of the American Academy of Political and Social Science*, 589(1), 41-62.
- Petrosino, A., Turpin-Petrosino, C., Hollis-Peel, M. E., & Lavenberg, J. G. (2013). 'Scared Straight' and other juvenile awareness programs for preventing juvenile delinquency. *Cochrane database of systematic reviews*(4).
- Petteruti, A., Schindler, M., & Ziedenberg, J. (2014). Sticker Shock: Calculating the Full Price Tag for Youth Incarceration. Retrieved from <http://www.justicepolicy.org/research/8477>
- Petteruti, A., Walsh, N., & Velázquez, T. (2009). The costs of confinement: Why good juvenile justice policies make good fiscal sense. *Justice Policy Institute*.
- Poulin, F., & Boivin, M. (2000). Reactive and proactive aggression: evidence of a two-factor model. *Psychological assessment*, 12(2), 115.
- Price, J. M., & Dodge, K. A. (1989). Reactive and proactive aggression in childhood: Relations to peer status and social context dimensions. *Journal of abnormal child psychology*, 17(4), 455-471.
- Ray, J. V., Frick, P. J., Thornton, L. C., Steinberg, L., & Cauffman, E. (2016). Positive and negative item wording and its influence on the assessment of callous-unemotional traits. *Psychological assessment*, 28(4), 394.
- Renshaw, T. L. (2016). Psychometric properties of the Social and Emotional Health Survey with a small sample of academically at-risk adolescents. *Journal of Psychoeducational Assessment*, 34(5), 487-495.
- Resnick, M. D., Bearman, P. S., Blum, R. W., Bauman, K. E., Harris, K. M., Jones, J., . . . Shew, M. (1997). Protecting adolescents from harm: findings from the National Longitudinal Study on Adolescent Health. *Jama*, 278(10), 823-832.

- Rhodes, J. (2002). *New directions for youth development: A critical view of youth mentoring*. San Francisco: Jossey-Bass.
- Riley, A., & Weiss, M. R. (2015). Summary Report: Spring Evaluation. *Girls on the Run International*.
- Roth, J. L., & Brooks-Gunn, J. (2016). Evaluating youth development programs: Progress and promise. *Applied Developmental Science, 20*(3), 188-202.
- Royse, D. (1998). Mentoring high-risk minority youth: Evaluation of the Brothers Project. *Adolescence, 33*(129), 145-159.
- Rutter, M. (1979). Protective factors in children's responses to stress and disadvantage. *Annals of the Academy of Medicine, Singapore, 8*(3), 324-338.
- Sapotichne, B. (2015). *Fearful temperament moderates the effect of harsh parenting on early childhood problem behaviors within dangerous neighborhoods: A multilevel analysis* (Doctoral dissertation). University of New Orleans Theses and Dissertations. 2106. <https://scholarworks.uno.edu/td/2106>
- Schwab-Stone, M. E., Ayers, T. S., Kaspro, W., Voyce, C., Barone, C., Shriver, T., & Weissberg, R. P. (1995). No safe haven: A study of violence exposure in an urban community. *Journal of the American Academy of Child & Adolescent Psychiatry, 34*(10), 1343-1352.
- Sherman, L. W., Gottfredson, D. C., MacKenzie, D. L., Eck, J., Reuter, P., & Bushway, S. (1997). *Preventing crime: What works, what doesn't, what's promising: A report to the United States Congress*: US Department of Justice, Office of Justice Programs Washington, DC.
- Shields, A., & Cicchetti, D. (1998). Reactive aggression among maltreated children: The contributions of attention and emotion dysregulation. *Journal of clinical child psychology, 27*(4), 381-395.
- Simons, R. L., Wu, C. I., Conger, R. D., & Lorenz, F. O. (1994). Two routes to delinquency: Differences between early and late starters in the impact of parenting and deviant peers. *Criminology, 32*(2), 247-276.
- Simpkins, S., Little, P., Weiss, H., & Simpkins-Chaput, S. (2004). Understanding and measuring attendance in out-of-school programs. *Issues and Opportunities in Out-of-School Time Evaluation Briefs, 7*.
- Slade, E. P., & Wissow, L. S. (2007). The influence of childhood maltreatment on adolescents' academic performance. *Economics of education review, 26*(5), 604-614.
- Slicker, E. K., & Palmer, D. J. (1993). Mentoring at-risk high school students: Evaluation of a school-based program. *The School Counselor, 40*(5), 327-334.
- Smith, D. K., Leve, L. D., & Chamberlain, P. (2011). Preventing internalizing and externalizing problems in girls in foster care as they enter middle school: Impact of an intervention. *Prevention Science, 12*(3), 269-277.
- Stimmel, M. A., Cruise, K. R., Ford, J. D., & Weiss, R. A. (2014). Trauma exposure, posttraumatic stress disorder symptomatology, and aggression in male juvenile offenders. *Psychological trauma: theory, research, practice, and policy, 6*(2), 184.
- Thombs, D. L. (2000). A retrospective study of DARE: Substantive effects not detected in undergraduates. *Journal of Alcohol and Drug Education, 46*(1), 27.
- Tierney, J. P., & Grossman, J. B. (1995). *Making a Difference: An Impact Study of Big Brothers/Big Sisters 1988-1995*. Retrieved from

- Timmons-Mitchell, J., Bender, M. B., Kishna, M. A., & Mitchell, C. C. (2006). An independent effectiveness trial of multisystemic therapy with juvenile justice youth. *Journal of Clinical Child & Adolescent Psychology, 35*(2), 227-236.
- Tolan, P., Henry, D., Schoeny, M., Bass, A., Lovegrove, P., & Nichols, E. (2013). Mentoring Interventions to Affect Juvenile Delinquency and Associated Problems: A Systematic Review. *Campbell Systematic Reviews 2013: 10. Campbell Collaboration.*
- Tremblay, R. E., Masse, B., Perron, D., LeBlanc, M., Schwartzman, A. E., & Ledingham, J. E. (1992). Early disruptive behavior, poor school achievement, delinquent behavior, and delinquent personality: longitudinal analyses. *Journal of consulting and clinical psychology, 60*(1), 64.
- Tull, M. T., Barrett, H. M., McMillan, E. S., & Roemer, L. (2007). A preliminary investigation of the relationship between emotion regulation difficulties and posttraumatic stress symptoms. *Behavior Therapy, 38*(3), 303-313.
- Valois, R. F., MacDonald, J. M., Bretous, L., Fischer, M. A., & Drane, J. W. (2002). Risk factors and behaviors associated with adolescent violence and aggression. *American journal of health behavior, 26*(6), 454-464.
- van der Veen-Mulders, L., Nauta, M. H., Timmerman, M. E., van den Hoofdakker, B. J., & Hoekstra, P. J. (2017). Predictors of discrepancies between fathers and mothers in rating behaviors of preschool children with and without ADHD. *European child & adolescent psychiatry, 26*(3), 365-376.
- Vitaro, F., Brendgen, M., & Tremblay, R. E. (2002). Reactively and proactively aggressive children: Antecedent and subsequent characteristics. *Journal of Child Psychology and Psychiatry, 43*(4), 495-505.
- Wandersman, A., & Florin, P. (2003). Community interventions and effective prevention. *American psychologist, 58*(6-7), 441.
- Wasserman, G. A., Keenan, K., Tremblay, R. E., Coie, J. D., Herrenkohl, T. I., Loeber, R., & Petechuck, D. (2003). Risk and Protective Factors of Child Delinquency. OJJDP Child Delinquency Bulletin Series.
- West, S. L., & O'Neal, K. K. (2004). Project DARE outcome effectiveness revisited. *American journal of public health, 94*(6), 1027-1029.
- White, S. F., Cruise, K. R., & Frick, P. J. (2009). Differential correlates to self-report and parent-report of callous-unemotional traits in a sample of juvenile sexual offenders. *Behavioral sciences & the law, 27*(6), 910-928.
- Widom, C. S. (1989). The cycle of violence. *Science, 244*(4901), 160.
- Williford, A. P., Calkins, S. D., & Keane, S. P. (2007). Predicting change in parenting stress across early childhood: Child and maternal factors. *Journal of abnormal child psychology, 35*(2), 251-263.
- Wu, P., Hoven, C. W., Bird, H. R., Moore, R. E., Cohen, P., Alegria, M., . . . Lichtman, J. H. (1999). Depressive and disruptive disorders and mental health service utilization in children and adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry, 38*(9), 1081-1090.
- Zamora, D. (2005). Levels of academic achievement and further delinquency among detained youth. *The Southwest Journal of Criminal Justice, 2*(1), 42-53.

Vita

Molly Miller is from Overland Park, KS. In 2012, she received her B.S. in Psychology and in 2016 earned her M.S. in Applied Psychology, both from the University of New Orleans. Her research focuses on prevention and reduction of antisocial behavior among youth in the juvenile justice system and within the community.