The impact of two universal randomized first- and second-grade classroom interventions on young adult suicide ideation and attempt[*]

Holly C. Wilcox[a,*], **Sheppard G. Kellam**[b], **C. Hendricks Brown**[c], **Jeanne Poduska**[b], **Nicholas S. Ialongo**[d], **Wei Wang**[c], and **James C. Anthony**[e]

a *Department of Psychiatry & Behavioral Sciences, Johns Hopkins School of Medicine, 600 North Wolfe Street/CMSC 346, Baltimore, MD 21287, United States*

b *American Institutes for Research, 921 E. Fort Avenue, Suite 225, Baltimore, MD 21230, United States*

c *Department of Epidemiology and Biostatistics, College of Public Health, University of South Florida, 13201 Bruce B Downs Boulvard, Tampa, FL 33612, United States*

d *Johns Hopkins University, Bloomberg School of Public Health 624 N. Broadway, 8th Floor, Baltimore, MD 21205, United States*

e *Department of Epidemiology, College of Human Medicine, Michigan State University B601 West Fee Hall, East Lansing, MI 48824, United States*

Abstract

Objective—This paper reports the impact of two first- and second-grade classroom based universal preventive interventions on the risk of Suicide Ideation (SI) and Suicide Attempts (SA) by young adulthood. The Good Behavior Game (GBG) was directed at socializing children for the student role and reducing aggressive, disruptive behavior. Mastery Learning (ML) was aimed at improving academic achievement. Both were implemented by the teacher.

Methods—The design was epidemiologically based, with randomization at the school and classroom levels and balancing of children across classrooms. The trial involved a cohort of first-grade children in 19 schools and 41 classrooms with intervention at first and second grades. A replication was implemented with the next cohort of first grade children with the same teachers but with little mentoring or monitoring.

Results—In the first cohort, there was consistent and robust GBG-associated reduction of risk for suicide ideation by age 19–21 years compared to youths in standard setting (control) classrooms regardless of any type of covariate adjustment. A GBG-associated reduced risk for suicide attempt was found, though in some covariate-adjusted models the effect was not statistically robust. No statistically significant impact on these outcomes was found for ML. The impact of the GBG on suicide ideation and attempts was greatly reduced in the replication trial involving the second cohort.

Conclusions—A universal preventive intervention directed at socializing children and classroom behavior management to reduce aggressive, disruptive behavior may delay or prevent onset of suicide ideation and attempts. The GBG must be implemented with precision and continuing support of teachers.

Keywords

Suicide ideation; Suicide attempts; Adolescence; Developmental epidemiology; Universal prevention programs; Good behavior game

1. Introduction

Suicide and suicide-related behavior have been studied in epidemiology for more than 150 years. Among the earliest contributions were those made by William Farr, the early 19th century father of biostatistics, who sought to test whether suicide risk varied with successful adaptation or achievement in relation to education (Farr, 2000). Later in the 19th century, Emile Durkheim's sociological theory linked higher rates of suicide to the absence of shared social values and norms (anomie) in the general population and lower rates of suicide with the opposite—greater social integration and shared values (Durkheim, 1897). In the present study we extend that research into the domain of experimental and developmental epidemiology at the school and classroom levels. The main aim of this paper is to estimate the impact of two universal classroom-based interventions on risk of Suicide Ideation (SI) and/or Suicide Attempts (SA), referred to here as *suicidality*. We report on evidence from a randomized controlled trial of two preventive interventions implemented in the first and second grade classrooms of elementary schools with follow-up assessments of suicidality when the participants were in young adulthood. One intervention, the Good Behavior Game (GBG), was directed at producing shared values, norms, and proper student behavior within the first- and second-grade classroom. The second, Mastery Learning (ML), was directed at learning to read, a basic task of social adaptation in first and second grades.

Suicide ideation is considered by some to be the first step toward suicide (Gili-Planas et al., 2001); it is one of the strongest predictors of a future suicide attempt in adolescents and is related to risk of completed suicide (Brent et al., 1993; Lewinsohn et al., 1996; Reinherz et al., 1995). The current public health significance of suicidal thoughts and behaviors is large in the United States and globally (Murray and Lopez, 1996; United States Public Health Service, 1999). Prevalence estimates derived from recent community studies of adolescents indicate that up to 30% have experienced suicide ideation (Fergusson and Lynskey, 1995; Ialongo et al., 2002; Juon and Ensminger, 1997; Lewinsohn et al., 1996; Reinherz et al., 1995; Velez and Cohen, 1988). In the United States alone, approximately two million adolescents attempt suicide each year, resulting in 700,000 emergency room visits, and death certifications in 2005 for 15–24 year-olds indicated that over 4000 had committed suicide (McIntosh, 2008; Shaffer

and Pfeffer, 2001). Prevention of suicide has been declared a national priority (United States Public Health Service, 1999), but there is little definitive evidence on the efficacy or effectiveness of suicide prevention programs among children and adolescents (Burns and Patton, 2000; Goldsmith et al., 2002; Gould et al., 2003).

Over the past three decades, evidence from developmental epidemiological studies has consistently identified specific antecedents measured as early as first grade in the prediction of later mental and behavioral disorders during the middle school years and beyond (Cairns et al., 1989; Ensminger et al., 1983; Farrington, 1995; Hawkins et al., 1988, 1992; Kellam et al., 1983; Reid, 1993; Reid and Eddy, 1997). Randomized preventive intervention trials conducted by our group and others indicate that school-based universal interventions (i.e., those addressing all children, not merely those indicated to be at higher risk) can have beneficial effects on aggressive, disruptive behavior and achievement (Dolan et al., 1993; Ialongo et al., 1999, 2001; Kellam et al., 1994a; Reid et al., 1999), off-task behavior (Brown, 1993b), depressive symptoms in first grade (Kellam et al., 1994b), and delayed onset or reduced risk of tobacco smoking (Kellam and Anthony, 1998).

The work reported in this paper is grounded in *life course/social field theory* (Kellam et al., 1975), which is built on the observation that one or more main social fields are critically important in each stage of life. Within each of these social fields there are defined social task demands to which an individual must respond, a process called *social adaptation*. The adequacy of an individual's response to these demands, termed *social adaptational status* (SAS), is rated by *natural raters* in each social field, such as parents in the family, teachers in the classroom, significant peers in the peer group, or, later in the life course, by spouses in the intimate social field or supervisors in the work social field. In keeping with Cicchetti and Schneider-Rosen (1984), the theory posits that success in mastering social task demands specific to one stage of development and in one social field will lead to an increase in later successes in the same and other social fields. Life course/social field theory also suggests that SAS often is reciprocally related to *psychological well-being* (PWB), the internal psychological status of an individual in regard to constructs such as self-esteem, psychiatric symptoms or disorders, and neurobiological or neuropsychological conditions. PWB may be an antecedent and/or a consequence of SAS, and vice versa.

The social task demands of the first grade classroom have been studied and reported frequently by our research group (e.g., Kellam et al., 1994a,b). In general, these social task demands consist of socializing appropriately with other children and the teacher, obeying classroom rules, concentrating and attending, and learning academic subjects. The two interventions used in this study were directed at improving social adaptation to these demands; therefore, we hypothesized that decreasing aggressive, disruptive behavior in classrooms would lead to a reduction in risk of outcomes linked to this antecedent (Kellam et al., 1994a), and that improving classroom achievement would improve PWB, particularly depressive symptoms and disorders in vulnerable children (Kellam et al., 1994b).

As the course of development unfolds in relation to successes or failures at social adaptational tasks in classrooms, in peer groups, and the community, other mediating or moderating issues can stem from as well as further influence the course of development and the impact of the GBG or ML. In community studies of adolescents, the use of alcohol and illegal drugs has been identified as an influential factor contributing to observed increases in the rate of adolescent suicide (Brent et al., 1988; Fowler et al., 1986). The prevalence of substance use disorders (alcohol and other drugs) in psychological autopsy and other studies has varied between 37% and 66% (Brent et al., 1988; Fowler et al., 1986; Runeson, 1989; Shaffer et al., 1988; Shafii et al., 1988). Depression, substance abuse, and aggressive, disruptive behavior have been found to distinguish suicide attempters from non-attempters in community and

clinical studies of adolescents and adults (Brent et al., 1993; Garrison et al., 1993; Kessler et al., 1999; Lewinsohn et al., 1996; Petronis et al., 1990). Drug and alcohol abuse has been associated with greater frequency of suicide attempts, more serious attempts in terms of lethality and intent, and increased levels of suicide ideation (Crumley, 1990; Lewinsohn et al., 1996). Garrison et al. (1993) found that the relationships between alcohol and illegal drug use and suicidal behaviors were most pronounced with the reported use of the more potentially dangerous or 'harder' drugs (e.g., cocaine) but remained even when the substance of interest was nicotine. The role of early and continuing drug use may be an important element in the developmental psychopathology of suicidality and will be considered in this report. Other potential mediators of this association could be constructs often thought to be associated with early drug and alcohol use, such as aggressive, disruptive behavior in later childhood or early adolescence, Conduct Disorder symptoms, academic self-competence, self-derogation, parental monitoring, or drug using or deviant peers.

1.1. The good behavior game

The GBG was developed by Barrish et al. (1969) and was selected for this trial because of its efficacy in short-term non-experimental or quasi-experimental designs, and because it was acceptable to the participating school leadership and community. The GBG is a classroom team-based behavior management strategy that promotes good behavior by rewarding teams that do not exceed maladaptive behavior standards as set by the teacher. The goal of the GBG is to create an integrated classroom social system that is supportive of all children being able to learn with little aggressive, disruptive behavior. The methods involve helping teachers to define unacceptable behaviors clearly and to socialize children with regulation of teammate's behavior through a process of team contingent reinforcement and mutual self-interest. In this trial, children in the GBG classrooms were assigned to one of three heterogeneous teams containing equal numbers of boys and girls, equal numbers of aggressive, disruptive children, and equal numbers of shy, socially isolated children. At the start of the game session, the teacher described and posted basic classroom rules of student behavior. All teams could "win" during a particular game period, with the criterion for winning the game being less than four infractions of acceptable student behavior.

Evidence to date has been supportive of the GBG reducing behavioral problems in the primary and middle school years, as well as reducing risk of tobacco smoking in early adolescence, mainly for boys (Kellam et al., 1994a; Kellam and Anthony, 1998). The GBG intervention trial also has shown long-term impact on drug abuse and dependence and alcohol abuse and dependence disorders as well as antisocial personality disorder and regular tobacco smoking, especially among males who had higher levels of aggressive, disruptive behavior in first grade (Kellam et al., this issue).

1.2. Mastery learning

ML is a teaching strategy with demonstrated effectiveness in improving achievement and the underlying theory and research posit that under appropriate instructional conditions virtually all students can learn most of what they are taught (Block and Burns, 1976; Bloom, 1976; Dolan, 1986; Guskey, 1997). Prior to and throughout this trial, all first-grade teachers in the Baltimore City Public School System were expected to teach reading using a ML curriculum. Our ML intervention teachers received a more precise, enhanced reading curriculum that they implemented in the context of a system-wide ML curriculum (Dolan et al., 1993; Kellam et al., 1991). This enhanced ML curriculum was used by both first- and second-grade teachers. Short-term ML benefits were found in reading achievement (Brown, 1993a; Dolan et al., 1993), and children with depressive symptoms in the fall of first grade who gained in achievement showed reduced depressive symptoms by the end of first grade (Kellam et al., 1994b).

1.3. Hypothesized mechanisms for ML and GBG impact on suicidality

The occurrence of suicidality among teenagers has been linked to the targets of both the ML and GBG interventions. Low reading achievement at school entry and other academic problems have been linked with suicidality, usually via an association with conduct problems (Beautrais et al., 1996, 1997; Bennett et al., 2003; Lewinsohn et al., 1994;) or depressive symptoms and disorder (Andrews and Lewinsohn, 1992; Beautrais et al., 1996; Brent et al., 1993; Esposito and Clum, 2002; Shaffer et al., 1996). Depressive symptoms and disorder are also considered strong predictors of suicidality in youth and young adults (Andrews and Lewinsohn, 1992; Beautrais et al., 1998; Fergusson and Lynskey, 1995; Kessler et al., 1999; Petronis et al., 1990; Reinherz et al., 1995).

The proximal target of the GBG, aggressive, disruptive behavior has also been reported as a risk factor for suicidality (Brent et al., 1993) and a number of externalizing behaviors. Such behavior outcomes are developmentally linked to aggressive, disruptive behavior and include impulsivity (Stanley et al., 1986), conduct disorder (Andrews and Lewinsohn, 1992; Shaffer et al., 1996), antisocial behavior (Beautrais et al., 1996; Fergusson and Lynskey, 1995; Shafii et al., 1985), juvenile justice involvement (Beautrais et al., 1997; Fergusson and Lynskey, 1995) and alcohol and other drug use disorders (Beautrais et al., 1996; Brent et al., 1993; Shaffer et al., 1996).

Our previously reported findings support a conclusion that the ML intervention affects its intended proximal targets. Specifically, ML showed a short-term impact on learning and mastery, particularly for those who began school with high levels of depressive symptoms (Kellam et al., 1994b). By lowering depressive symptoms, ML could hypothetically also lessen suicidality through improved mastery of the student role. Furthermore, the ML intervention appeared to have somewhat higher overall achievement results for children who began more depressed in first grade, and improvement in achievement was associated with a lessening of self-reported depressed feelings (Kellam et al., 1994b). We therefore hypothesized that ML would reduce suicidality because of the role that mastery plays in self-esteem, depression, and suicidal thoughts (Battle, 1990; Brookover, 1965; Covington, 1989; Holly, 1987).

The GBG impact was also supported by previous findings. Students in the GBG classrooms showed reduced aggressive, disruptive behavior during elementary and middle school as well as reduced drug and alcohol involvement through young adulthood (Kellam et al., this issue). Because the GBG intervenes early in the child's life before classroom maladaptive aggressive, disruptive behavior becomes intransigent, we hypothesize that the GBG should reduce the course of aggressive, disruptive behavior and many of its consequences, including suicidality. Several reasons for this continuity in aggressive, disruptive behavior are made clear in the Patterson–Reid model of conduct disorder (Reid et al., 2002). In the absence of an intervention to reduce aggressive, disruptive behavior, aggressive, disruptive children may have trouble learning academic subjects, may remain disliked by teachers and their peers, and often limit their friends to other deviant peers (Reid et al., 2002). Over time, such children are at increased risk of failing in school and engaging in drug use, and impulsive and antisocial behavior (Kellam et al., 1983, 1994a, 1998, this issue). All of these processes and conditions place youths at higher risk for suicide (Gould et al., 2003).

In prior papers we have reported that the GBG intervention had a long-term impact into young adulthood on outcomes such as drug and alcohol abuse and dependence disorders, tobacco use, antisocial personality disorder and violent and criminal behavior, and service use, particularly among males who were at higher levels of first grade aggressive, disruptive behavior, with lesser impact among females (Kellam and Anthony, 1998; Kellam et al., this issue; Petras et al., this issue; Poduska et al., this issue). We therefore hypothesize that the largest GBG impact will be seen in highly aggressive, disruptive males. We have also reported that females show

greater developmental continuity in internalizing symptoms (Kellam et al., 1983). Suicidality could have internalizing and/or externalizing pathways into young adulthood; there could also be different developmental pathways for females and males. In addition, ML might alter internalizing while the GBG might impact externalizing pathways.

2. Methods

The research design was that of an epidemiologically based randomized field trial, applied to a defined population of children. The study population consisted of all youths who started first grade in 41 classrooms in 19 elementary schools of the Baltimore City Public School System during two successive academic years: 1985–1986 for Cohort 1 first graders and 1986–1987 for Cohort 2 first graders. In this paper emphasis is placed on the first cohort where the highest level of mentoring and monitoring for fidelity was conducted. The 19 schools were located within five urban areas which varied in socioeconomic status from very poor to moderate income, and in degree of racial segregation. Within each area, three or four schools were sorted into homogenous sets, and selected for the trial (Kellam et al., 1991). All of the entering first graders in the schools became eligible participants in the study. Of the 1196 pupils in Cohort 1, 66% were African-American and almost all the remainder were non-Hispanic Whites. Parents, and children as they grew older, could refuse to participate in assessments at any time. Written consent was obtained from parents for childhood participation for those assessments beyond standard school assessment procedures. Consent was obtained from each young adult at the time of the young adult interviews. During the first year of recruitment for each cohort, only 5% declined to participate or sign consent forms (Kellam et al., this issue); participation to the young adult years is described below.

Within each intervention school, all incoming first grade children were assigned to classrooms so as to balance classrooms within school on gender, kindergarten experience, and academic and behavioral performance. Following this, classrooms/teachers were assigned at random to intervention condition, that is, in schools testing ML, they were assigned to either ML or standard program (internal ML control). The same procedure was followed within the GBG schools, yielding GBG classrooms and internal GBG control classrooms. Within the control schools, all youth were assigned in a balanced fashion to classrooms that served as external controls. In a few instances, the study team observed imbalance after assignment, so a small number of children were reassigned immediately to induce balance on entry to first grade. The ML and GBG interventions lasted for 2 years. To maintain this design, GBG and ML students moved into second grade in their same classroom units. Similarly, standard program (control) students were kept together in the control condition through the end of second grade. Transfers from one intervention to the other or from control to either intervention were rare. We have analyzed the data based on original assignment (i.e., in an 'intent to treat' design). After the 2-year intervention period ended, students were assigned by the school to classrooms via usual and customary procedures of the school system without regard to prior intervention status. The same design was used for the second cohort and teachers delivered the same intervention as they did for the first cohort.

Both ML teachers and GBG teachers each received an initial 40 h of training supporting their respective interventions. Standard program (control) teachers received a comparable time of in-service training but in areas not affecting the experimental intervention conditions. Care was taken to provide comparable incentives and reinforcements to all teachers in the five conditions.

Over 15 years after school entry, between 1999 and 2002, the research team assessed 1918 (83%) in the two young adult follow-up interviews, including 154 participants found within incarceration facilities (who were interviewed with additional consent approvals). Four percent

were traced but refused to be interviewed; another 6% were traced, but for logistical reasons were not interviewed (e.g., out of state with no telephone number, military postings overseas); 5.7% were not found despite extensive tracing efforts. National Death Index searches through April 2003 found that 41 pupils had died (1.8%), including one confirmed suicide.

The key response variables in this study were the occurrence of suicidality as assessed by face-to-face interview during young adulthood. Other young adult information was also obtained from a separate phone interview (see Kellam et al., this issue). The study protocols for the two follow-up assessments were reviewed and approved by the Institutional Review Boards for protection of human subjects in research at the Johns Hopkins Bloomberg School of Public Health and the phone interview by the American Institutes for Research as well.

The components of suicidality, suicide ideation and suicide attempt, were analyzed separately. Age at first suicide ideation and age at first attempt were assessed via standardized questions embedded within the follow-up interview's multi-item assessment of major depressive disorder based upon an adapted version of the NIMH Diagnostic Interview Schedule (DIS) (Robins et al., 1981, 1994). Suicide ideation was measured in the face-to-face survey by asking, "Have you ever felt so low you thought about committing suicide?" Youths who answered 'yes' have been designated as cases of suicide ideation, and were asked to report the age of first occurrence of ideation. Suicide attempt data were collected with the question, "Have you ever attempted suicide?"

Other covariates under study were selected on the basis of prior theory and research on suicide-related behaviors (e.g., see Brent, 1995; Garrison et al., 1991, 1993; Ialongo et al., 2002; Juon and Ensminger, 1997; Lewinsohn et al., 1994, 1996). A standardized teacher rating of aggressive, disruptive behavior was taken in the fall of first grade, which provided an assessment of early aggressive, disruptive behavior or rule-breaking as described elsewhere (e.g., Kellam et al., 1975; Werthamer-Larsson et al., 1991). The child's scaled self-reports of depressive symptoms (Child Depression Index-CDI) and anxious symptoms (Revised Children's Manifest Anxiety Scale, RCMAS) in fall of first grade were also used (Ialongo et al., 1994, 2001; Kellam et al., 1994b). Data for each youth's gender, age, race/ethnicity, and family financial status (i.e., eligibility for free or reduced lunch) were abstracted from the school system's administrative database. Familial suicidality and psychiatric disturbance was assessed retrospectively during a follow-up assessment conducted during early young adulthood via telephone.

2.1. Measurement of potential mediators

Early onset alcohol, tobacco, marijuana, and inhalant use (use before age 16) was measured in yearly assessments during 1989–1994 in school via interviewer administered self-report methods. The children were first interviewed about their use of alcohol, tobacco, marijuana, and inhalants in the spring of 1989, when they were in third or fourth grades, and thereafter through 1994 in eighth or ninth grades. Standardized questions on age at first use were as follows: "How old were you when you first smoked tobacco?"; "Not counting sips with your parent's permission, how old were you the first time you drank beer, wine, wine coolers, liquor or any other drink with alcohol in it?"; "How old were you the first time you smoked marijuana?"; "(Not counting cocaine,) How old were you the first time you sniffed something to get high or for some other feeling like that?" Additional information on these standardized assessments of youthful drug involvement can be obtained from prior publications of this research group (e.g., see Chilcoat et al., 1995; Chilcoat and Anthony, 1996; Kellam and Anthony, 1998).

In 1989, when the children were in third or fourth grade, an adapted version of the Harter Scholastic Competence subscale (Harter, 1985) was used to assess academic self-competence.

Self-derogation was measured by a slightly modified version of the Kaplan–Pokorny Scale (Kaplan and Pokorny, 1969). Also in 1989, overt and covert antisocial behavior scales were adapted from the Patterson–Capaldi OSLC measure of Overt Antisocial behavior (Capaldi and Patterson, 1989). The Baltimore Conduct Problems and Delinquency Scale was used in 1991. This self-report measure of delinquent and antisocial behavior is an adaptation of the measure developed by Elliott and Huizinga for the National Survey of Delinquency and Drug Use (Elliott et al., 1985). Drug using peer involvement was assessed in 1989 by an eight item yes/no scale that assessed the frequency of peers who smoke tobacco, drink alcohol, smoke marijuana, use crack, frequency of peers who say it is a bad idea to use drugs, and peers who would drop people as friends if they started using drugs. Peer involvement in rule-breaking behavior was assessed in 1989 by a five item scale that assessed the frequency of peers who cheat on tests, ruin or damage property belonging to someone else, steal something worth less than five dollars, threaten to hit someone without a reason and suggest that the participant do something against the law. This scale was a modified version of that described by Capaldi and Patterson (1989). Youths were asked in forced choice format to indicate how often their peers have engaged in the various antisocial behaviors. Parental monitoring was assessed in 1989 (Capaldi and Patterson, 1989). The 10-item modified scale was originally from the Structured Interview of Parent Management Skills and Practices—Youth Version (SIPMSP, Capaldi and Patterson, 1989). Major Depressive Episode was assessed using the Major Depressive Disorder section of the CIDI, Version 2.1, which was modeled after the National Institute of Mental Health-Diagnostic Interview Schedule (DIS).

Survival analysis methods were used to estimate the relative hazard of suicide ideation and attempt for both ML and the GBG, followed by covariate adjustment for gender, race, baseline aggressive, disruptive behavior, clinical features of depression and anxiety, parental suicidality and psychiatric disturbance. All survival analyses measured the time to first suicide ideation (or attempt) since entry into first grade; the two participants who reported age of onset of suicide ideation before age six were excluded from analyses. For those who did not report suicidality, the censoring time was age at last assessment. Separate Kaplan–Meier curves for each intervention condition and gender are presented for those in Cohort 1 along with log-rank test statistics and 95% confidence intervals (CI).

To take into account the fact that time was measured in years, we used Discrete-Time Survival Analyses (DTSA) with covariates, as suggested by Cox (1972), and more recently by Willet and Singer (1993) and Singer and Willet (1994). To account for clustering of students within initial elementary schools and within urban areas, we stratified on school. This stratified analysis accommodates data interdependencies by conditioning on school and thus allows direct comparison of each of the two active interventions with their respective internal controls. We have found that school variation was high at baseline across a number of measures. Once we condition on school many variables show little variation at the classroom level because children were balanced across classrooms (Brown, 1993b; Kellam et al., this issue). Thus, balancing classes within schools is an efficient way to control for school and community characteristics in impact analyses (e.g., see Breslow and Day, 1980; Brown and Liao, 1999; Schlesselman, 1982). This approach is especially useful when suicidality might depend on socially shared determinants that are not easily quantifiable (e.g., widely shared religious beliefs and local area norms about suicide-related behavior).

3. Results

3.1. Characteristics of the population and the epidemiology of suicidality

Table 1 presents characteristics of the Cohort 1 sample at baseline and follow-up including the cumulative incidence of suicide ideation and suicide attempt (SA) among young adult participants, as well as unadjusted relative risk estimates for gender, race, free lunch status,

and randomized intervention assignment. The mean age at the face-to-face follow-up interview was 22 years (range 20–24 years old) for Cohort 1 and 21 years (range 19–23 years old) for Cohort 2.

3.1.1. Suicide ideation—In Cohort 1, 123 of 858 participants (14%) assessed in young adulthood had experienced SI, which was twice as common (RR = 2.2) among non-Hispanic White youths as within the other racial/ethnic group, nearly all of which were African American ($p < 0.001$). It was also 1.3 times more common among participants who had not qualified for free or subsidized lunch versus those who had ($p = 0.041$). Those individuals assigned to the GBG intervention were half as likely to have experienced SI, as compared to those in the control classrooms (internal GBG, internal ML, and external controls, RR = 0.5, $p = 0.024$). The ideation value for the ML intervention also was smaller than that for controls, but was not statistically significant ($p = 0.211$).

In the Cohort 2 sample, 95 of 837 participants (11%) assessed in young adulthood had experienced SI, which was almost three times as common (RR = 2.7) among non-Hispanic White youths as within the other racial/ethnic group, nearly all of which were African American ($p < 0.001$). SI was 1.5 times more common among participants who had not qualified for free or subsidized lunch versus those who had ($p = 0.001$). Approximately, 9% of those who had received the GBG intervention had experienced SI compared to 12% of those in the control classrooms, but the relative risk estimate did not reach statistical significance (RR = 0.8, $p = 0.408$). The ML impact on SI was also not statistically significant (RR = 1.1, $p = 0.592$). Full Cohort 2 analyses for this and other outcomes can be found in the supplementary material on the journal's website.

3.1.2. Suicide attempt—In Cohort 1, 94 of 858 participants (11%) assessed in young adulthood had made an SA. The cumulative incidence of SA was nearly twice as high among females as males (RR = 1.9, $p = 0.007$). The rate was also 70% higher among non-Hispanic Whites compared to within the other racial/ethnic group ($p = 0.014$). The crude SA cumulative incidence estimates did not vary by free or reduced lunch status. They were slightly lower in both intervention groups compared to controls ($p > 0.15$ for both).

In the Cohort 2 sample, 61 of 837 participants (7%) assessed in young adulthood had experienced SA, which was two times as common (RR = 2.0) among non-Hispanic White youths as within the other racial/ethnic group, nearly all of which were African American ($p = 0.010$). Those individuals assigned to the GBG intervention were not less likely to have experienced SA compared to controls, as the relative risk estimate did not reach statistical significance (RR = 1.0, $p = 0.965$). The SA relative risk estimate for the ML intervention was also not statistically significant (RR = 1.1, $p = 0.844$). In both Cohorts, one-third of those who had attempted suicide, ideation was not present suggesting that the SA was an impulsive act without anticipatory ideation.

3.2. Kaplan–Meier life table results

We compared onset of SI by intervention condition among females in Cohort 1 (see Fig. 1). There was markedly lower occurrence of SI in the GBG condition compared to controls after age 16. By young adulthood, the cumulative incidence of SI was 19% for females in the control condition versus about 9% for those in the GBG condition. Fig. 1b shows a much smaller difference in SI incidence for ML females versus internal ML controls. Fig. 2 shows plotted cumulative incidence values for SA among females. Here, the patterns were similar to SI with a reduction in risk for the GBG but not for ML. About 20% of controls had made at least one SA by young adulthood versus 10% for the GBG.

Parallel results for SI and SA among males in Cohort 1 are presented separately in Figs. 3 and 4. Fig. 3a shows that by age 15, the SI incidence for the GBG group decreased relative to control values. By young adulthood, approximately 24% of males who were in control classrooms had experienced SI versus 11% of GBG males. Just as for females, the ML impact on SI was null for males; approximately 15% of boys in ML classrooms had SI versus 13% in the control classrooms. Ten percent of GBG males had made an SA by age 22 versus 18% of internal GBG control males. There were no differences between ML and controls for males' SA risk. For full Cohort 2 analyses please refer to the supplementary material on the journal website.

3.3. GBG and ML impact on suicidal ideation

Table 2 presents relative odds of the GBG impact based on Discrete-Time Survival Analysis for SI for Cohort 1. Here, given SI rarity, the relative odds (RO) approximates relative risk (RR). Four models are presented. Model 1 is unadjusted for covariates; Model 2 adjusts for gender, race, and baseline levels of aggressive, disruptive behavior, depression, and anxiety as measured in fall of first grade; Model 3 adjusts for covariates in Model 2 as well as caregiver suicidality and mental illness that were obtained retrospectively at the young adult interviews; Model 4 is Model 2 plus terms to capture a hypothesized interaction of design and baseline aggressive, disruptive behavior, depression, and anxiety. Given no gender difference in the GBG impact ($p = 0.828$) for SI and SA, Tables 2 and 3 include data on both males and females.

The overall summary relative risk estimate for occurrence of SI indicates an inverse (protective) association with assignment to the GBG intervention (RR = 0.4; 95% CI = 0.2, 0.8; $p = 0.008$), as well as a consistent and robust inverse association of the GBG intervention with SI across the series of covariate-adjusted models (RR = 0.4–0.5, see Table 2). The intervention impact did not vary by baseline mental health indicators; no design by baseline interaction terms in Model 4 were statistically significant (all $p > 0.05$). A model was then fit for SI that included covariates with relative risk estimates that were greater than 2.0—namely GBG and caregiver suicide ideation, both of which showed strong and consistent associations with SI across the series of models. The best fitting model for SI in Cohort 1 included terms for GBG (RR = 0.5; 95% CI 0.2, 0.9; $p = 0.015$) and caregiver SI (RR = 4.3; 95% CI 2.5, 7.2; $p < 0.001$). In contrast to findings for the GBG intervention, relative risk estimates for ML based upon Discrete-Time Survival Analysis for SI for Cohort 1 were not statistically significant (all $p > 0.05$, data not shown in table).

Gender-stratified results revealed that the estimates for females were very similar to the overall estimates. For Cohort 1 females, the unadjusted association showed a relative risk of 0.40 (95% CI 0.2, 0.9; $p = 0.042$, data not shown in a table) and adjustment for baseline mental health and demographic characteristics did not change the estimate appreciably. The GBG impact was statistically significant in all models. In the series of models, caregiver SI was the only covariate that was statistically significant (RR = 3.4; 95% CI 1.1, 10.7; $p = 0.037$). In the Cohort 1 male analyses, the GBG had a marginal but non-robust association with SI with or without adjustment for baseline covariates (RR = 0.4, 95% CI 0.2, 1.1, $p = 0.067$; RR = 0.4, 95% CI 0.2, 1.1, $p = 0.083$). The association weakened in models adjusting for the series of covariates. The effects for males on SI were thus slightly smaller, with larger p-values than those for females. The relative risk estimate for caregiver SI was even stronger for males than females.

Relative risk estimates for ML were all non-significant, but in the hypothesized direction (RR < 1.0) for Cohort 1 females and in the opposite direction (RR > 1.0) for Cohort 1 males. For females in Cohort 1, one interaction term (depression × ML) was significant in Model 4 (RR = 5.7, 95% CI 1.2, 28.6; $p = 0.033$). ML-assigned females in the top quartile for depression in fall of first grade were found to experience increased risk for subsequent SI. Of note is the fact that neither baseline aggressive, disruptive behavior nor depressive symptoms by themselves predicted suicide ideation in males or females. There was no evidence of an interaction effect

of baseline aggressive, disruptive behavior or baseline depressive symptoms with the GBG intervention condition.

Using the same series of models, the association between the GBG and SI was examined in Cohort 2. The impact of the GBG on SI in Cohort 2 was in the same direction as Cohort 1 but not statistically significant compared to internal GBG controls (RR = 0.6–0.8, $p \geq 0.2$). In the series of adjusted models, SI was four times as common among non-Hispanic White youths as within other racial/ethnic groups (RR = 3.9–4.3, $p \leq 0.006$), nearly all of which were African American. Those with caregiver SI were three and a half times more likely to report SI themselves (RR = 3.5; 95% CI 1.6, 7.4; $p = 0.001$). The impact of ML on SI was null compared to internal ML controls (RR = 1.2–1.5, $p \geq 0.2$).

Whereas the Cohort 1 female results were consistent with a robust GBG-associated protection against SI, this was not the case for the Cohort 2 replication. For females, the GBG impact on SI risk was modest (RR = 0.4–0.8; $p \geq 0.1$) as compared to female internal GBG controls. White non-Hispanic ethnicity and caregiver SI were the only two covariates that were associated with SI by young adulthood (RR = 4.6; 95% CI 1.2, 17.6; $p = 0.027$; RR = 3.3; 95% CI 1.1, 9.9; $p = 0.032$, respectively from the best fitting model). Estimates for the GBG impact on SI for Cohort 2 males as compared to male internal GBG controls were inverse but weaker, and caregiver SI was the only covariate that was associated with SI (RR = 3.6, 95% CI 1.0, 13.1; $p = 0.049$). For both females and males, results for ML were all non-significant as compared to respective female and male internal ML controls, but the results were in the opposite direction than the GBG impact (RR > 1.0).

3.4. GBG and ML impact on suicide attempt

Corresponding estimates for GBG and risk of SA for Cohort 1 are presented in Table 3. The overall summary relative risk estimate for occurrence of SA is consistent with an inverse (protective) association with assignment to the GBG intervention (RR = 0.5, 95% CI 0.3, 0.9; $p = 0.041$). Subsequently, adjusting for an array of covariates, this association of the GBG retained strength, but for some models lost statistical significance (see Table 3, Model 2 GBG RR = 0.5, 95% CI 0.3–1.0, $p = 0.065$; Model 3 GBG RR = 0.6, 95% CI 0.3, 1.2; $p = 0.130$). Covariate terms for baseline by intervention interactions yielded no sign of variation in intervention impact as a function of baseline characteristics. Female gender, baseline depression, and caregiver mental illness retained statistically significance. The best fitting model for SA in Cohort 1 included terms for GBG (RR = 0.5; 95% CI 0.3, 0.9; $p = 0.031$) and female gender (RR = 2.0; 95% CI 1.3, 3.1; $p = 0.002$).

We also explored subgroup variation via stratified analyses, with initial attention to gender differences (data not shown in a table). There was evidence of a GBG-associated SI and SA risk differences in females, but the interaction between the GBG and gender in the Discrete-Time Survival Analysis models were not statistically significant by conventional standards ($p > 0.05$). All Cohort 1 female models of GBG impact on SA were in the predicted direction. In analyses for Cohort 1 males, all estimates were in the predicted direction (RR < 1.0) except the model that adjusted for caregiver variables. No ML impact estimates for SA were statistically significant although all estimates were in the hypothesized direction (data not in a table). There was evidence that females in the highest aggressive, disruptive quartile in the fall of first grade were at increased risk for suicide attempt (RR = 7.7; 95% CI 1.4, 42.5; $p = 0.020$). There was no other evidence of an interaction effect of baseline aggressive, disruptive behavior or baseline depressive symptoms with the GBG intervention condition.

For full Cohort 2 analyses please refer to the supplementary material on the journal website. In Cohort 2 gender stratified analyses, neither the GBG nor ML were significantly associated with SA in these models. White non-Hispanic ethnicity for males and caregiver suicide threats

for females were the only covariates in the series of models that were statistically significant (RR = 5.7; 95% CI 1.2, 27.0; $p = 0.028$ and RR = 20.2; 95% CI 2.9, 142.5; $p = 0.003$, respectively). There were estimation problems in some of the Cohort 2 male analyses, as the data were too sparse for computation.

Analyses were conducted individually for the six GBG schools to see if the impact of the GBG intervention on SI or SA varied by school in Cohort 1. The impact of the GBG on both SI and SA was particularly strong in one urban school that was characterized as 99% African American and 93% of students eligible for free lunch; estimates of the GBG impact in the other schools were consistently favoring the GBG but of somewhat smaller magnitude.

3.5. Examination of mediation

The examination of mediation in our study addresses the extent to which change in one or more mediating variables account for observed program effects. The following variables were tested as possible mediators in the association between the GBG and suicide ideation and attempt: alcohol use before age 16, tobacco before age 16, marijuana use before age 16, inhalant use before age 16, Conduct Disorder symptoms, Academic Self-Competence, Self-Derogation, Parental Monitoring, Overt Aggression, Covert Aggression, Drug Using Peers, Deviant Peers and depressive episode that occurred prior to suicide ideation or attempt. For both continuous and categorical mediators, the product of coefficient method was used to test the significance of a mediation effect (MacKinnon et al., 2002). This approach treats mediation as the product of two regression coefficients, the regression of the mediator on treatment conditions, and the partial regression coefficient of the outcome on the mediator in Discrete-Time Survival Analyses, adjusted for the treatment conditions. We tested both suicide attempt and suicide ideation and the analyses were conducted separately for males and females. According to the method used, none of the variables studied showed evidence of potential mediation.

4. Discussion

This paper reports on one of a few but growing set of epidemiologically based randomized prevention trials to study children's school achievement, behavior, and psychological well-being at the time of entry into first grade and to conduct assessments in young adulthood. In this classroom-based randomized trial the results support the hypothesis that first graders assigned to GBG classrooms experienced subsequent lower incidence of suicidality through childhood, adolescence, and into young adulthood compared to internal GBG controls. They reported half the lifetime rates of ideation and attempts compared to their matched controls. For ideation, the beneficial GBG effect was consistent regardless of whether baseline covariates were included. The GBG effect on attempts was less definitive once we controlled for gender and baseline depressive symptoms. This beneficial impact of the GBG was consistent for Cohort 1 males and females, who show similar relative risks. There was no consistent finding of positive relations between first grade aggressive, disruptive behavior or depressive symptoms and suicidality. Caregiver suicide ideation had a strong and consistent association with offspring's suicide ideation, especially among males, and retrospective reports on caregiver mental illness and suicide threats were also related to offspring suicide attempts by young adulthood.

This study is one of the few universal population-based randomized trials to evaluate impact on suicidality (Brown et al., 2007). Most other trials have been based on selected high-risk samples. Recently, however, there have been trials on a curriculum that teaches self-awareness of depression (Aseltine and DeMartino, 2004) and on a gatekeeper program to detect those who are suicidal (Brown et al., 2006; Wyman et al., 2008). Although suicidality in young people represents a critical public health issue, the relatively low cumulative incidence of suicidal behaviors necessitates relatively large sample sizes or special randomized designs

(Brown et al., 2006; see also Brown and Faraone, 2004). This study gains power from a substantial sample size and a long follow-up period.

A few study limitations merit attention. The results seen in this trial can only be directly generalized to the given population and time period from which the sample came. This issue of generalizablity can be addressed only through replication in other school districts with similar and different social, ethnic, and economic characteristics. A trial of the GBG with a similar cohort of children has been conducted in the Netherlands and the findings tend to provide evidence of effectiveness of the GBG (van Lier et al., 2005). The beneficial GBG findings for suicidality in Cohort 1 were not replicated in Cohort 2, although most Cohort 2 findings were in a consistent direction. The GBG was implemented in the second cohort with less precision because sufficient mentoring and monitoring procedures were not in place, resulting in two possible consequences: a diminished impact and/or a shift in impact (see Kellam et al., this issue). There is also a hypothetical statistical explanation for these diminished findings in the second cohort. Classroom variability and heterogeneity across the three control conditions were markedly higher in the second cohort compared to the first; hence, statistical power was reduced.

Another limitation involves measurement. Suicide ideation necessarily requires some degree of self-report assessment because many people with suicidal ideation do not come to clinical attention and this aspect of psychological well-being requires interpersonal communication if it is to be known by others. The self-report character of the study data and the constrained number of survey items on suicidality are both limitations, as was the reliance upon retrospective assessment of age of onset during young adulthood. Also, since the caregiver data were collected retrospectively from the youths, it is possible that the statistical model for estimating the GBG impact was mis-specified when terms for caregiver mental illness and suicidality were included. These terms might well attenuate the impact of the GBG through biased underreporting or partial mediation since these events could occur after intervention. Relationships found with caregiver reports might be explained by the fact that suicidal youth would be more likely to know about their parents' psychiatric histories than youth who did not show suicidal behavior. Finally, the minimal relationship that we found between baseline aggressive, disruptive behavior and depressive symptoms on the one hand and suicidality on the other, in contrast to the known relation between these factors, could be partially explained by the long time interval from elementary school to adulthood.

Our analytical method in this multi-level study relied on an analysis that stratified by school, which has the advantage of holding fixed contextual characteristics related to neighborhood and school that were not possible to include explicitly in our analyses. While classroom characteristics were not formally taken into account in this analysis, our conditioning on school should address this factor since within schools there was only one control classroom and at most two intervention classrooms. Thus, estimates from this analysis, comparing suicidality within schools should be very similar to estimates from random effects models that formally take classroom variation into account. Such random effects models also appropriately take into account the group-based randomization in this study.

It was difficult to examine some of the assumptions underlying our analysis. We addressed these interdependencies via a survival analysis approach conditioning on the schools within which classrooms were nested. However, multilevel models also may prove to be useful to quantify intervention effects in the presence of interdependencies (e.g., the alternating logistic regressions). In addition, we examined some mediational hypotheses, but this trial had limited capacity to carry out more detailed mediational models across the life course to examine specific mechanisms leading to reduced suicidality.

The GBG has been found to have short- and long-term impact on externalizing behaviors among males in Cohort 1 and to a more limited extent in Cohort 2 especially among aggressive, disruptive first grade males (Kellam et al., 1994a, this issue). The long-term impact of the GBG on alcohol and drug abuse and dependence disorders (Kellam et al., this issue), mental health service use (Poduska et al., this issue), and Antisocial Personality Disorder in young adulthood (Petras et al., this issue) was consistently strong particularly among higher risk males, whereas impact on anxiety and depressive disorders has been found non-significant in intent-to-treat analyses (Kellam et al., this issue). This report on suicidality showed a stronger and more consistent GBG impact in Cohort 1, similar to the findings in our other studies in this issue; although this suicidality study showed similar impact among both females and males. The only other outcome that held for both genders was alcohol abuse and dependency disorders, which for females was not related to their early aggressive, disruptive behavior. The reasons for this difference in impact of the GBG may be related to alcohol use disorders and suicidality not being as anti-social as the drug and ASPD/violence impacts of the GBG also reported in this issue. Clearly, the developmental epidemiology of these issues and gender differences in such studies are a vital next stage in understanding impact and increasing it for the GBG and other interventions.

The results are generally consistent with life course/social field theory regarding the importance of early mastery of social task demands in the classroom in promoting later successful social adaptation not only in school but also in other main social fields. This work resonates with that of Durkheim (1897) in that when teachers cannot manage their classrooms there is an absence of clear norms (anomie). The GBG intervention was aimed at setting the social contextual norms, with a goal of pupils' mastery of the student role. The GBG consists of: (1) establishing and posting classroom rules, teaching children the meaning of the rules and providing examples of breaking the rules; (2) the teacher assigning students to teams within the classroom setting consisting of equal numbers of girls and boys and heterogeneous with regard to aggressive, disruptive behavior. The teacher thereby establishes both classroom norms of behavior and group membership of each child; these are reinforced by the children in each team influencing each other to follow the teacher established rules. We hypothesize that this deliberate teacher led classroom structure around norms and membership was how the GBG might have influenced both the behavior and the psychological well-being of the children. van Lier et al. (2005) found that, in classrooms with the GBG where the group structure was established by the teachers, children who started with more aggressive, disruptive behavior not only improved, but they associated with children later on who were better behaving. Children are strongly affected by their social contexts, and the teacher's ability to improve the social adaptational process may indeed promote mastery in children, thereby leading to improved sense of self and self-esteem, and thereby raising confidence in one's competence to deal with future demands in school and beyond (Harris et al., 1990; Kellam et al., 1994a, 1998). In addition, the GBG reduced aggressive, disruptive and off-task behavior and presumably therefore allowed teachers more time/opportunity to teach (Brown, 1993b).

The Good Behavior Game and Mastery Learning are "universal interventions" received by all children in all classrooms and are not focused upon youths at higher risk. As such, the interventions are generally economical and do not require much extra teacher or student burden; they can be implemented by teachers during the regular school day. The absence of findings for ML suggests that the GBG effect was specific to the GBG, and cannot be attributed to added attention in the GBG classrooms. However, the GBG intervention must be carried out with precision. The results were more modest in a second cohort of first graders with the same teachers, but without the support of 40 h of mentoring and monitoring that these same teachers had received during the prior year with the first cohort. Based on extensive analyses of alternative explanations, we hypothesize that the absence of retraining and monitoring of program implementation during the second cohort resulted in more modest impact. In addition,

the prevalence of suicidality was lower in the second cohort as compared to the first cohort, reflecting, in part, that Cohort 1 was an average of 1 year older.

The results point toward the need for training programs for new teachers and in-service training for more experienced ones. Teachers are not as often trained in classroom behavior management as strongly as these results suggest they should be. The results also underline the vital importance of the first grade classroom as a stage setting the course for children's life course in school and beyond (Kellam et al., this issue). The implications of these findings are that the first grade classroom can have a critical impact on the developmental course of psychopathology. While this seems an apparent concept, the strength and clarity of the impact of the precisely directed early GBG intervention emphasizes the validity of early interventions in this critical social field.

Acknowledgements

References

Andrews JA, Lewinsohn PM. Suicidal attempts among older adolescents: prevalence and co-occurrence with psychiatric disorders. J Am Acad Child Adolesc Psychiatry 1992;31:655–662. [PubMed: 1644728]

Aseltine RH, DeMartino R. An outcome evaluation of the SOS suicide prevention program. Am J Publ Health 2004;94:446–451.

Barrish H, Saunders M, Wolf M. Good Behavior Game: effects of individual contingencies for group consequences on disruptive behavior in a classroom. J Appl Behav Anal 1969;2:119–124. [PubMed: 16795208]

Battle, J. Self-Esteem: The New Revolution. James Battle & Assoc; Alberta, Canada: 1990.

Beautrais AL, Joyce PR, Mulder RT. Risk factors for serious suicide attempts among youths aged 13 through 24 years. J Am Acad Child Adolesc Psychiatry 1996;35:1174–1182. [PubMed: 8824061]

Beautrais AL, Joyce PR, Mulder RT. Precipitating factors and life events in serious suicide attempts among youths aged 13 through 24 years. J Am Acad Child Adolesc Psychiatry 1997;36:1543–1551. [PubMed: 9394939]

Beautrais AL, Joyce PR, Mulder RT. Psychiatric illness in a New Zealand sample of young people making serious suicide attempts. N Z Med J 1998;111:44–48. [PubMed: 9539914]

Bennett KJ, Brown KS, Boyle M, Racine Y, Offord D. Does low reading achievement at school entry cause conduct problems? Soc Sci Med 2003;56:2443–2448. [PubMed: 12742607]

Block, J.; Burns, R. Mastery learning. In: Shulman, L., editor. Review of Research in Education. Vol. 4. F.E. Peacock; Itasca, IL: 1976. p. 3-49.

Bloom, BS. Human Characteristics and School Learning. McGraw-Hill; New York: 1976.

Brent DA. Risk factors for adolescent suicide and suicidal behavior: mental and substance abuse disorders, family environmental factors, and life stress. Suicide Life Threat Behav 1995;25:52–63. [PubMed: 8553429]

Brent DA, Johnson B, Bartle S, Bridge J, Rather C, Matta J, Connolly J, Constantine D. Personality disorder, tendency to impulsive violence, and suicidal behavior in adolescents. J Am Acad Child Adolesc Psychiatry 1993;32:69–75. [PubMed: 8428886]

Brent DA, Perper JA, Kolko DJ, Zelenak JP. The psychological autopsy: methodological considerations for the study of adolescent suicide. J Am Acad Child Adolesc Psychiatry 1988;27:362–366. [PubMed: 3379020]

Breslow NE, Day NE. Statistical methods in cancer research, vol. 1 The analysis of case–control studies IARC. Sci Publ 1980;32:5–338.

Brookover, WB. Office of Research and Public Information. Michigan State University; East Lansing, MI: 1965. Self-Concept of Ability and School Achievement.

Brown CH. Analyzing preventive trials with generalized additive models. Am J Community Psychol 1993a;21:635–664. [PubMed: 8192125]

Brown CH. Statistical methods for preventive trials in mental health. Stat Med 1993b;12:289–300. [PubMed: 7681214]

Brown, CH.; Faraone, SV. Prevention of schizophrenia and psychotic behavior: definitions and methodologic issues. In: Stone, WS.; Faraone, SV.; Tsuang, MT., editors. Early Clinical Intervention and Prevention in Schizophrenia. Humana Press; Totowa, NJ: 2004. p. 255-284.

Brown CH, Liao J. Principles for designing randomized preventive trials in mental health: an emerging developmental epidemiology paradigm. Am J Community Psychol 1999;27:673–710. [PubMed: 10676544]

Brown CH, Wyman PA, Brinales JM, Gibbons RD. The role of randomized trials in testing interventions for the prevention of youth suicide. Int Rev Psychiatry 2007;19:617–631. [PubMed: 18092240]

Brown CH, Wyman PA, Guo J, Peña J. Dynamic wait-listed designs for randomized trials: new designs for prevention of youth suicide. Clin Trials 2006;3:259–271. [PubMed: 16895043]

Burns JM, Patton GC. Preventive interventions for youth suicide: a risk factor-based approach. Aust N Z J Psychiatry 2000;34:388–407. [PubMed: 10881963]

Cairns RB, Cairns BD, Neckerman HJ. Early school dropout: configurations and determinants. Child Dev 1989;60:1437–1452. [PubMed: 2612252]

Capaldi, DM.; Patterson, GR. Psychometric Properties of Fourteen Latent Constructs from the Oregon Youth Study. Springer-Verlag; New York: 1989.

Chilcoat HD, Anthony JC. Impact of parent monitoring on initiation of drug use through late childhood. J Am Acad Child Adolesc Psychiatry 1996;35:91–100. [PubMed: 8567618]

Chilcoat HD, Dishion TJ, Anthony JC. Parent monitoring and the incidence of drug sampling in urban elementary school children. Am J Epidemiol 1995;141:25–31. [PubMed: 7801962]

Cicchetti, D.; Schneider-Rosen, K. Toward a transactional model of childhood depression. In: Cicchetti, D.; Schneider-Rosen, K., editors. Childhood Depression: A Developmental Perspective. Jossey-Bass; San Francisco: 1984. p. 5-28.

Covington, M. Self-esteem and failure in school. In: Mecca, A.; Smelser, NJ.; Vasconcellos, J., editors. The Social Importance of Self-Esteem. UC Press; Berkeley, CA: 1989.

Cox DR. Regression models and life-tables. J R Stat Soc B 1972;34:187–220.

Crumley FE. Substance abuse and adolescent suicidal behavior. JAMA 1990;263:3051–3056. [PubMed: 2188024]

Dolan L. Mastery learning as a preventive strategy. Outcomes 1986;5:20–27.

Dolan LJ, Kellam SG, Brown CH, Werthamer-Larsson L, Rebok GW, Mayer LS, Laudolff J, Turkkan J, Ford C, Wheeler L. The short-term impact of two classroom-based preventive interventions on aggressive and shy behaviors and poor achievement. J Appl Dev Psychol 1993;14:317–345.

Durkheim, E. Suicide: A Study of Sociology. Spaulding, JA.; Simpson, G., editors. Free Press; Glencoe, IL: 1897. Translated, 1951

Elliott, DS.; Huizinga, D.; Ageton, SS. Explaining Delinquency and Drug Use. Sage Publications; Beverly Hills, CA: 1985.

Ensminger, ME.; Kellam, SG.; Rubin, BR. School and family origins of delinquency: comparisons by sex. In: Van Dusen, KT.; Mednick, SA., editors. Prospective Studies of Crime and Delinquency. Kluwer-Nijhoff Publishing; Boston: 1983. p. 73-97.

Esposito C, Clum GA. Psychiatric symptoms and their relation to suicidal ideation in a high risk adolescent community sample. J Am Acad Child Adolesc Psychiatry 2002;41:44–51. [PubMed: 11800204]

Farr W. Vital statistics: memorial volume of selections from the reports and writings 1885. Bull World Health Org 2000;78:88–95. [PubMed: 10733295]

Farrington DP. The development of offending and antisocial behavior from childhood: key findings from the Cambridge Study in Delinquent Development. J Child Psychol Psychiatry 1995;360:929–964.

Fergusson DM, Lynskey MT. Childhood circumstances, adolescent adjustment, and suicide attempts in a New Zealand birth cohort. J Am Acad Child Adolesc Psychiatry 1995;34:612–622. [PubMed: 7775356]

Fowler RC, Rich CL, Young D. San Diego Suicide Study II: substance abuse in young cases. Arch Gen Psychiatry 1986;43:962–965. [PubMed: 3753161]

Garrison CZ, Addy CL, Jackson KL, McKeown RE, Waller JL. The CES-D as a screen for depression and other psychiatric disorders in adolescents. J Am Acad Child Adolesc Psychiatry 1991;30:636–641. [PubMed: 1890099]

Garrison CZ, McKeown RE, Valois RF, Vincent ML. Aggression, substance use, and suicidal behaviors in high school students. Am J Public Health 1993;83:179–184. [PubMed: 8427319]

Gili-Planas M, Roca-Bennasar M, Ferrer-Perez V, Bernardo-Arroyo M. Suicidal ideation, psychiatric disorder, and medical illness in a community epidemiological study. Suicide Life Threat Behav 2001;31:207–213. [PubMed: 11459253]

Goldsmith, SK.; Pellmar, TC.; Kleinman, AM.; Bunney, WE. Reducing Suicide: A National Imperative. National Academies Press; Washington, DC: 2002. p. 1-516.

Gould MS, Greenberg T, Velting DM, Shaffer D. Youth suicide risk and preventive interventions: a review of the past 10 years. J Am Acad Child Adolesc Psychiatry 2003;42:386–405. [PubMed: 12649626]

Guskey, TR. Implementing Mastery Learning. Vol. 2. Wadsworth Publishing Co; Belmont, CA: 1997.

Harris TO, Brown GW, Bifulco AT. Depression and situational helplessness/mastery in a sample selected to study childhood parental loss. J Affect Disord 1990;20:27–41. [PubMed: 2147188]

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

Hawkins JD, Catalano RF, Miller JY. Risk and protective factors for alcohol and other drug problems in adolescence and early adulthood: implications for substance abuse prevention. Psychol Bull 1992;112:64–105. [PubMed: 1529040]

Hawkins JD, Doueck HJ, Lishner DM. Changing teaching practices in mainstream classrooms to improve bonding and behavior of low achievers. Am Educ Res J 1988;25:31–50.

Holly, W. Oregon School Study Council. University of Oregon; Eugene, OR: 1987. Self-Esteem: Does It Contribute to Students' Academic Success?.

Ialongo NS, Edelsohn G, Kellam SG. A further look at the prognostic power of young children's reports of depressed mood and feelings. Child Dev 2001;72:736–747. [PubMed: 11405579]

Ialongo N, Edelsohn G, Werthamer-Larsson L, Crockett L, Kellam S. The significance of self-reported anxious symptoms in first-grade children. J Abnorm Child Psychol 1994;22:441–455. [PubMed: 7963077]

Ialongo N, McCreary BK, Pearson JL, Koenig AL, Wagner BM, Schmidt NB, Poduska J, Kellam SG. Suicidal behavior among urban. African American young adults Suicide Life. Threat Behav 2002;32:256–271.

Ialongo NS, Werthamer L, Kellam SG, Brown CH, Wang S, Lin Y. Proximal impact of two first-grade preventive interventions on the early risk behaviors for later substance abuse, depression, and antisocial behavior. Am J Community Psychol 1999;27:599–641. [PubMed: 10676542]

Juon HS, Ensminger ME. Childhood, adolescent, and young adult predictors of suicidal behaviors: a prospective study of African Americans. J Child Psychol Psychiatry 1997;38:553–563. [PubMed: 9255699]

Kaplan HB, Pokorny AD. Self-derogation and psychosocial adjustment. J Nerv Ment Dis 1969;149:421–434. [PubMed: 5347703]

Kellam SG, Anthony JC. Targeting early antecedents to prevent tobacco smoking: findings from an epidemiologically based randomized field trial. Am J Public Health 1998;88:1490–1495. [PubMed: 9772850]

Kellam, SG.; Branch, JD.; Agrawal, KC.; Ensminger, ME. Mental Health and Going to School: The Woodlawn Program of Assessment, Early Intervention, and Evaluation. University of Chicago Press; Chicago: 1975.

Wilcox HC. Effects of a universal classroom behavior management program in first and second grades on young adult behavioral, psychiatric, and social outcomes. Drug Alcohol Depend. this issue

Kellam, SG.; Brown, CH.; Rubin, BR.; Ensminger, ME. Paths leading to teenage psychiatric symptoms and substance abuse: developmental epidemiological studies in Woodlawn. In: Guze, SB.; Earls, FJ.; Barrett, JE., editors. Childhood Psychopathology and Development. Raven Press; New York: 1983. p. 17-51.

Kellam SG, Ling X, Merisca R, Brown CH, Ialongo N. The effect of the level of aggression in the first grade classroom on the course and malleability of aggressive behavior into middle school. Dev Psychopathol 1998;10:165–185. [PubMed: 9635220][See also Kellam SG, Ling X, Merisca R, Brown CH, Ialongo N. The effect of the level of aggression in the first grade classroom on the course and malleability of aggressive behavior into middle school: results of a developmental epidemiology-based prevention trial: Erratum. Dev Psychopathol 2000;12:107.].

Kellam SG, Rebok GW, Ialongo N, Mayer LS. The course and malleability of aggressive behavior from early first grade into middle school: results of a developmental epidemiologically-based preventive trial. J Child Psychol Psychiatry 1994a;35:259–282. [PubMed: 8188798]

Kellam SG, Rebok GW, Mayer LS, Ialongo N, Kalodner CR. Depressive symptoms over first grade and their response to a developmental epidemiologically based preventive trial aimed at improving achievement. Dev Psychopathol 1994b;6:463–481.

Kellam SG, Werthamer-Larsson L, Dolan LJ, Brown CH, Mayer LS, Rebok GW, Anthony JC, Laudolff J, Edelsohn G, Wheeler L. Developmental epidemiologically-based preventive trials: baseline modeling of early target behaviors and depressive symptoms. Am J Community Psychol 1991;19:563–584. [PubMed: 1755436]

Kessler RC, Borges G, Walters EE. Prevalence of and risk factors for lifetime suicide attempts in the National Comorbidity Survey. Arch Gen Psychiatry 1999;56:617–626. [PubMed: 10401507]

Lewinsohn PM, Rohde P, Seeley JR. Psychosocial risk factors for future adolescent suicide attempts. J Consult Clin Psychol 1994;62:297–305. [PubMed: 8201067]

Lewinsohn PM, Rohde P, Seeley JR. Adolescent suicidal ideation and attempts: prevalence, risk factors, and clinical implication. Clin Psychol Sci Prac 1996;3:25–46.

MacKinnon DP, Lockwood CM, Hoffman JM, West SG, Sheets V. A comparison of methods to test mediation and other intervening variable effects. Psychol Methods 2002;7:83–104. [PubMed: 11928892]

McIntosh, JL. American Association of Suicidology; 2008. U.S.A. Suicide: 2003 Official Final Data. http://www.suicidology.org/associations/1045/files/2003data.pdf

Murray, CJ.; Lopez, AD., editors. The Global Burden of Disease: A Comprehensive Assessment of Mortality and Disability from Diseases, Injuries, and Risk Factors in 1990 and Projected to 2020. Harvard School of Public Health. Harvard University Press; Cambridge, MA: 1996.

Petras H, Kellam SG, Brown CH, Muthén B, Ialongo NS, Poduska JM. Developmental epidemiological courses leading to antisocial personality disorder and violence and criminal behavior: effects by young adulthood of a universal preventive intervention in first-and second-grade classrooms. Drug Alcohol Depend. this issue

Petronis KR, Samuels JF, Moscicki EK, Anthony JC. An epidemiologic investigation of potential risk factors for suicide attempts. Soc Psychiatry Psychiatr Epidemiol 1990;25:193–199. [PubMed: 2399476]

Poduska J, Kellam S, Wang W, Brown CH, Ialongo N, Toyinbo P. Impact of the Good Behavior Game, a universal classroom-based behavior intervention, on young adult service use for problems with emotions, behavior, or drugs or alcohol. Drug Alcohol Depend. this issue

Reid JB. Prevention of conduct disorder before and after school entry: relating interventions to development findings. Dev Psychopathol 1993;5:243–262.

Reid, JB.; Eddy, JM. The prevention of antisocial behavior: some considerations in the search for effective interventions. In: Stoff, DM.; Breiling, J.; Maser, JD., editors. Handbook of Antisocial Behavior. John Wiley & Sons; New York: 1997. p. 343-356.

Reid JB, Eddy JM, Fetrow RA, Stoolmiller M. Description and immediate impacts of a preventive intervention for conduct problems. Am J Community Psychol 1999;27:483–517. [PubMed: 10573832]

Reid, JB.; Patterson, GR.; Snyder, JJ., editors. Antisocial Behavior in Children and Adolescents: A Developmental Analysis and the Oregon Model for Intervention. American Psychological Association; Washington, DC: 2002.

Reinherz HZ, Giaconia RM, Silverman AB, Friedman A, Pakiz B, Frost AK, Cohen E. Early psychosocial risks for adolescent suicidal ideation and attempts. J Am Acad Child Adolesc Psychiatry 1995;34:599–611. [PubMed: 7775355]

Robins LN, Helzer JE, Croughan J, Ratcliff KS. National Institute on Mental Health Diagnostic Interview Schedule: its history, characteristics, and validity. Arch Gen Psychiatry 1981;38:381–389. [PubMed: 6260053]

Robins, LN.; Helzer, JE.; Croughan, J.; Ratcliff, KS. National Institute on Mental Health Diagnostic Interview Schedule (DIS). In: Mezzich, JE.; Jorge, MR.; Salloum, IM., editors. Psychiatric Epidemiology: Assessment, Concepts and Method. The Johns Hopkins University Press; Baltimore, MD: 1994.

Runeson B. Mental disorder in youth suicide. DSM-III-R Axes I and II. Acta Psychiatr Scand 1989;79:490–497. [PubMed: 2750550]

Schlesselman, JJ. Case–Control Studies: Design, Conduct, Analysis. Oxford University Press; New York: 1982.

Shaffer D, Garland A, Gould M, Fisher P, Trautman P. Preventing teenage suicide: a critical review. J Am Acad Child Adolesc Psychiatry 1988;27:675–687. [PubMed: 3058676]

Shaffer D, Gould MS, Fisher P, Trautman P, Moreau D, Kleinman M, Flory M. Psychiatric diagnosis in child and adolescent suicide. Arch Gen Psychiatry 1996;53:339–348. [PubMed: 8634012]

Shaffer D, Pfeffer C. Practice parameters for the assessment and treatment of children and adolescents with suicidal behavior. J Am Acad Child Adolesc Psychiatry 2001;40:24S–51S. [PubMed: 11434483]

Shafii M, Carrigan S, Whittinghill JR, Derrick A. Psychological autopsy of completed suicide in children and adolescents. Am J Psychiatry 1985;142:1061–1064. [PubMed: 4025622]

Shafii M, Steltz-Lenarsky J, Derrick AM, Beckner C, Whittinghill JR. Comorbidity of mental disorders in the post-mortem diagnosis of completed suicide in children and adolescents. J Affect Disord 1988;15:227–233. [PubMed: 2975295]

Singer JD, Willet JB. Designing and analyzing studies on onset, cessation, and relapse: using survival analysis in drug abuse prevention research. NIDA Res Monogr 1994;142:196–263. [PubMed: 9243538]

Stanley M, Stanley B, Traskman-Bendz L, Mann JJ, Meyendorff E. Neurochemical findings in suicide completers and suicide attempters. Suicide Life Threat Behav 1986;16:286–300. [PubMed: 2428142]

United States Public Health Service. The Surgeon General's Call to Action to Prevent Suicide. United States Public Health Service; Washington, DC: 1999.

van Lier PAC, Vuijk P, Crijnen AAM. Understanding mechanisms of change in the development of antisocial behavior: the impact of a universal intervention. J Abnorm Child Psychol 2005;33:521–535. [PubMed: 16195948]

Velez CN, Cohen P. Suicidal behavior and ideation in a community sample of children: maternal and youth reports. J Am Acad Child Adolesc Psychiatry 1988;27:349–356. [PubMed: 3379018]

Werthamer-Larsson L, Kellam SG, Wheeler L. Effect of first grade classroom environment on child shy behavior, aggressive behavior, and concentration problems. Am J Community Psychol 1991;19:585–602. [PubMed: 1755437]

Willet JB, Singer JD. Investigating onset, cessation, relapse, and recovery: why you should, and how you can, use discrete-time survival analysis to examine event occurrence. J Consult Clin Psychol 1993;61:952–965. [PubMed: 8113496]

Wyman PA, Brown CH, Inman J, Cross W, Schmeelk-Cone K, Guo J, Pena JB. Randomized trial of a gatekeeper program for suicide prevention: 1-year impact on secondary school staff. J Consult Clin Psychol 2008;76(1):104–115. [PubMed: 18229988]

Fig. 1.
Estimated cumulative probability of reporting suicide ideation among females in Cohort 1 for GBG (a) and ML (b).

Fig. 2.
Estimated cumulative probability of reporting suicide attempt among females in Cohort 1 for GBG (a) and ML (b).
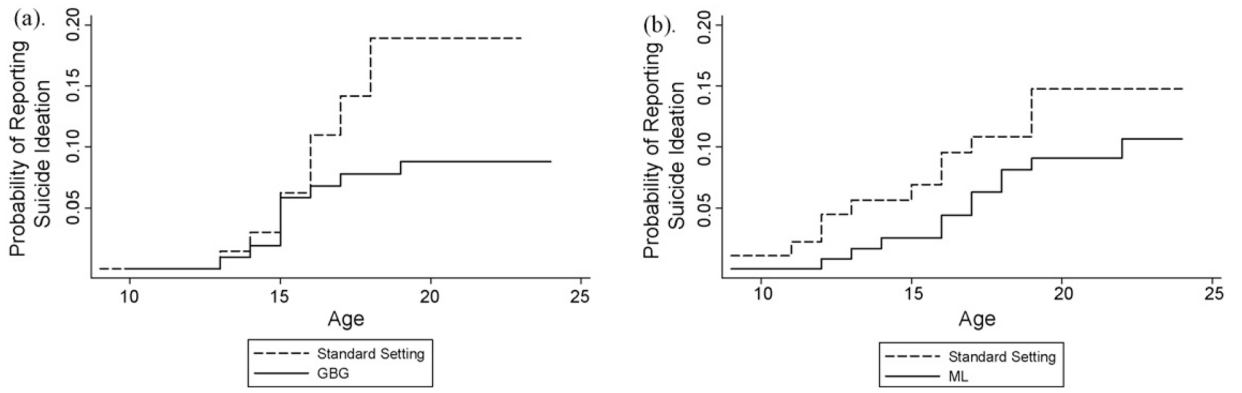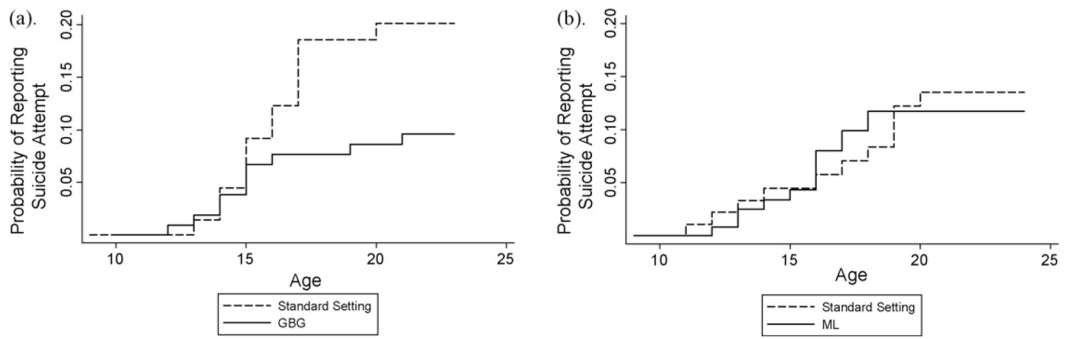
Fig. 3.
Estimated cumulative probability of reporting suicide ideation among males in Cohort 1 for
GBG (a) and ML (b).

Fig. 4.
Estimated cumulative probability of reporting suicide attempt among males in Cohort 1 for GBG (a) and ML (b).
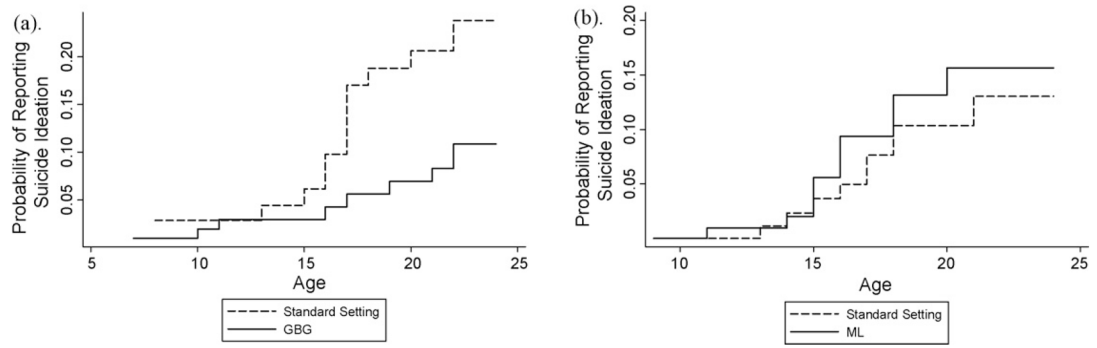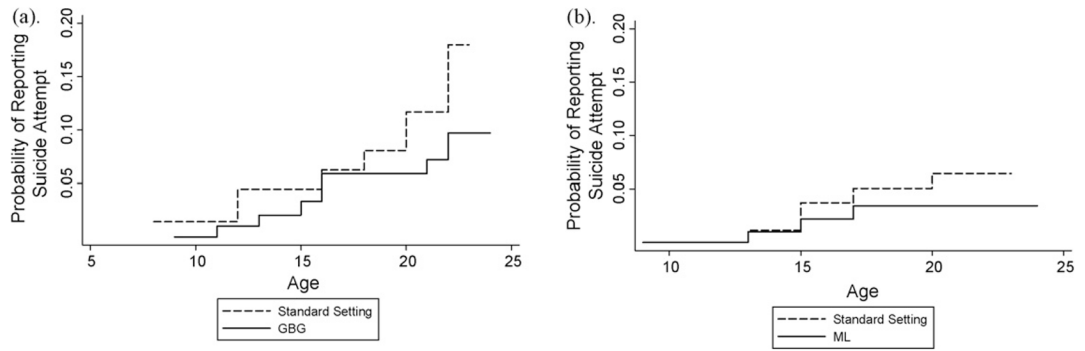
Table 1

Sample description and estimated risk of suicide ideation and attempt for Cohort 1

	At baseline	As young adults	Suicide ideation				Suicide attempt			
			Incident cases	Cumulative incidence	RR[a] 95% CI	p-Value	Incident cases	Cumulative incidence	RR[a] 95% CI	p-Value
Gender										
Male	587 (49%)	389 (45%)	58	0.15	1.0		30	0.08	1.0	
Female	609 (51%)	469 (55%)	65	0.14	0.9 (0.6–1.3)	0.662	64	0.14	1.9 (1.2–3.0)	0.007*
Race										
White (Non-Hispanic)	405 (34%)	233 (27%)	52	0.22	2.2		36	0.15	1.7	
Other racial/ethnic groups	791 (66%)	625 (73%)	71	0.11	1.0 (1.5–3.3)	<0.001*	58	0.09	1.0 (1.1–2.7)	0.014*
Free or partially subsidized lunch										
No	644 (55%)	431 (50%)	72	0.17	1.0		50	0.11	1.0	
Yes	548 (45%)	424 (50%)	50	0.12	0.7 (0.5–0.9)	0.041*	44	0.10	0.9 (0.6–1.4)	0.600
Intervention status										
Control	684 (57%)	492 (57%)	82	0.17	1.0		61	0.12	1.0	
ML	274 (23%)	188 (22%)	24	0.13	0.7 (0.4–1.2)	0.211	16	0.09	0.7 (0.4–1.2)	0.162
GBG	238 (20%)	178 (21%)	17	0.10	0.5 (0.3–0.9)	0.024*	17	0.09	0.7 (0.4–1.3)	0.292

[a] Relative risk estimates are not adjusted for covariates (gender, age, and race); adjusted estimates are presented in the text.

* $p < 0.05$.

Table 2
Impact of GBG on suicide ideation for males and females in Cohort 1

Predictor or covariate	Relative odds estimate (adjusted)	95% CI	p-Value
Model 1[a]			
GBG	0.4	0.2–0.8	0.008[*]
Model 2[b]			
GBG	0.4	0.2–0.8	0.011[*]
Female	1.0	0.7–1.5	0.828
White Non-Hispanic	1.4	0.7–3.1	0.376
Aggressive, disruptive behavior	0.9	0.6–1.4	0.624
Depressive symptoms	1.2	0.8–2.0	0.385
Anxious symptoms	1.2	0.5–2.7	0.647
Model 3[c]			
GBG	0.5	0.2–0.9	0.025[*]
Female	1.0	0.7–1.50	0.985
White Non-Hispanic	1.2	0.5–2.6	0.709
Aggressive, disruptive behavior	0.9	0.6–1.5	0.763
Depressive symptoms	1.2	0.7–2.0	0.442
Anxious symptoms	1.2	0.5–2.6	0.697
Caregiver depression	1.2	0.7–2.0	0.494
Caregiver ideation	2.5	1.2–5.5	0.018[*]
Caregiver suicide threat	1.7	0.7–4.1	0.269
Caregiver suicide attempt	1.0	0.4–2.7	0.961
Caregiver mentally ill	1.2	0.6–2.1	0.624
Caregiver mental health treatment	1.2	0.7–2.1	0.573
Model 4[d]			
GBG	0.4	0.2–0.9	0.044[*]
Female	1.0	0.7–1.5	0.842
White Non-Hispanic	1.5	0.7–3.2	0.323
Aggressive, disruptive behavior	1.0	0.6–1.8	0.980
Depressive symptoms	1.2	0.7–2.1	0.596
Anxious symptoms	1.6	0.6–4.1	0.362
Aggressive, disruptive behavior × GBG	0.8	0.3–2.9	0.824
Depressive symptoms × GBG	1.1	0.2–5.5	0.943
Anxious symptoms × GBG	1.4	0.1–14.1	0.776

[a]Unadjusted.

[b]Adjusted for sex, race, cohort, first grade aggressive, disruptive behavior, and symptoms.

[c]Adjusted for sex, race, cohort, first grade aggressive, disruptive behavior, and symptoms, as well as caregiver variables.

[d]Adjusted for sex, race, cohort, first grade aggressive, disruptive behavior, and symptoms, as well as interactions with GBG.

[*]$p < 0.05$.

Table 3

Impact of GBG on suicide attempt for males and females in Cohort 1

Predictor or covariate	Relative odds estimate (adjusted)	95% CI	*p*-Value
Model 1[a]			
GBG	0.5	0.3–0.9	0.041[*]
Model 2[b]			
GBG	0.5	0.3–1.0	0.065
Female	2.2	1.4–3.5	0.001[*]
White Non-Hispanic	2.0	0.8–5.0	0.142
Aggressive, disruptive behavior	1.4	0.9–2.4	0.143
Depressive symptoms	1.7	1.0–2.8	0.043[*]
Anxious symptoms	0.6	0.2–1.6	0.299
Model 3[c]			
GBG	0.6	0.3–1.2	0.130
Female	2.2	1.4–3.5	0.001[*]
White Non-Hispanic	1.8	0.7–4.6	0.243
Aggressive, disruptive behavior	1.6	0.9–2.7	0.070
Depressive symptoms	1.7	1.0–2.9	0.037[*]
Anxious symptoms	0.6	0.2–1.7	0.331
Caregiver depression	1.4	0.8–2.4	0.289
Caregiver ideation	1.7	0.7–4.0	0.250
Caregiver suicide threat	0.9	0.3–2.6	0.828
Caregiver suicide attempt	1.8	0.6–5.1	0.294
Caregiver mentally ill	2.2	1.1–4.2	0.021[*]
Caregiver mental health treatment	1.3	0.7–2.6	0.404
Model 4[d]			
GBG	0.3	0.1–0.7	0.008[*]
Female	2.3	1.4–3.6	0.001[*]
White Non-Hispanic	2.0	0.8–5.1	0.138
Aggressive, disruptive behavior	1.1	0.5–2.0	0.884
Depressive symptoms	1.5	0.8–2.8	0.199
Anxious symptoms	0.4	0.1–1.7	0.204
Aggressive, disruptive behavior × GBG	2.2	0.6–8.0	0.220
Depressive symptoms × GBG	1.2	0.2–6.2	0.869
Anxious symptoms × GBG	6.1	0.5–81.1	0.174

[a]Unadjusted.

[b]Adjusted for sex, race, cohort, first grade aggressive, disruptive behavior, and symptoms.

[c]Adjusted for sex, race, cohort, first grade aggressive, disruptive behavior, and symptoms, as well as caregiver variables.

[d]Adjusted for sex, race, cohort, first grade aggressive, disruptive behavior, and symptoms, as well as interactions with GBG.

[*]$p < 0.05$.