The Good Behavior Game and the Future of Prevention and Treatment

The Good Behavior Game (GBG), a universal classroom behavior management method, was tested in first- and second-grade classrooms in Baltimore beginning in the 1985–1986 school year. Followup at ages 19–21 found significantly lower rates of drug and alcohol use disorders, regular smoking, antisocial personality disorder, delinquency and incarceration for violent crimes, suicide ideation, and use of school-based services among students who had played the GBG. Several replications with shorter followup periods have provided similar early results. We discuss the role of the GBG and possibly other universal prevention programs in the design of more effective systems for promoting children's development and problem prevention and treatment services.

Sheppard G. Kellam, M.D.[1]
Amelia C. L. Mackenzie, B.S.[1]
C. Hendricks Brown, Ph.D.[2]
Jeanne M. Poduska, Sc.D.[3]
Wei Wang, Ph.D.[4]
Hanno Petras, Ph.D.[5]
Holly C. Wilcox, Ph.D.[6]

[1] Bloomberg School of Public Health
The Johns Hopkins University
Baltimore, Maryland

[2] Department of Epidemiology and Public Health
University of Miami
Miami, Florida

[3] American Institutes for Research
Washington, DC

[4] University of South Florida
Tampa, Florida

[5] JBS International
Bethesda, Maryland

[6] The Johns Hopkins University School of Medicine
Baltimore, Maryland

Drug, alcohol, and tobacco abuse and dependence disorders; antisocial personality disorder; violence; high-risk sexual behavior; and other disorders and problem behaviors impose huge personal, social, and economic costs on individuals, families, schools, and communities. The burden is borne also by institutions that treat or attempt to rehabilitate such problem behaviors and disorders.

Disruptive and aggressive behavior in classrooms as early as the first grade has repeatedly been identified as a risk factor for this spectrum of problems later in life (Kellam et al., 2008). The Good Behavior Game (GBG) is a classroom-wide, teacher-implemented intervention that aims to improve classroom behavior and introduce young children to the role of being a student and a member of the classroom community.

In 1985, in close partnership with the Baltimore City Public School System (BCPSS), we initiated a large-scale, developmental field trial of the GBG that was epidemiologically based and randomized. The trial was implemented in 41 first- and second-grade classrooms within 19 elementary schools with two consecutive cohorts of first graders. The results in young adulthood were reported in a supplemental issue of *Drug and Alcohol Dependence* in June 2008. Here we summarize the theoretical basis, design, and results of the trial, which together lead to three conclusions:

- Aggressive and disruptive behaviors in childhood play a causal role in a spectrum of social, behavioral, and psychiatric problems;
- Introducing the GBG in first- and second-grade classrooms reduces the risk of

A classroom playing the Good Behavior Game in Denver.

some of these problems later in the life course;

- The effectiveness of the GBG supports a role for universal prevention interventions in a redesigned system for child development and problem prevention and treatment.

We also briefly review the findings to date of ongoing replication trials and address the implications of this work for researchers, practitioners, advocates, and policymakers. We believe that the underlying theory, data, and analyses support the development of a newly designed human development services system that integrates prevention and treatment and is closely interrelated with schools and classrooms.

The GBG was developed to help teachers manage classrooms without having to respond on an individual basis each time a student disrupted class.

THE GOOD BEHAVIOR GAME

The GBG was developed to help teachers manage classrooms without having to respond on an individual basis each time a student disrupted class. As designed by University of Kansas researchers Harriet Barrish, Muriel Saunders, and Montrose Wolf, the GBG increases a teacher's precision and consistency in instructing elementary school students in appropriate classroom behavior. In documenting the effectiveness of the approach, an early observer noted reduced "talking out of turn" and "out of seat" behavior during times when the class played the GBG (Barrish, Saunders, and Wolf, 1969).

Our first-generation, large-scale randomized field trials of the GBG in Baltimore began in the 1985–1986 school year. By that time, the positive results reported by Barrish and colleagues had been replicated in more than 20 small observational, nonrandomized studies that showed short-term improvement in student classroom behavior.

How the Game Was Played

Teachers used a manual to ensure precision in the implementation of the GBG and to support fidelity over time and replicability in other trial sites. Early in the first-grade year, teachers displayed a large poster that listed the rules of proper student behavior—for example, sitting still, talking in turn, and paying attention. Toward the end of the first quarter of the school year, when classroom membership had stabilized, teachers divided their students into three teams that were balanced as to gender, aggressive and disruptive behavior, and shy or isolated behavior.

Initially, the GBG was played for designated periods of 10 minutes, three times a week. Each team was rewarded when all of its members behaved well during that interval, but not when the team had more than four rule infractions. In this way, the team's rewards were contingent on each member behaving well.

As the year continued, the GBG was played for increasing lengths of time and when students were working individually. In this way, the GBG facilitated learning without competing for instructional time. As the school year progressed, the rewards changed from tangible and immediate (e.g., stickers, erasers) to more abstract and deferred (e.g., gold stars, more time to do enjoyable activities).

Why the Teacher and Classroom?

The GBG treats the classroom as a community. The teacher is central to the GBG, because he or she sets the rules for becoming a successful student and member of the community and also determines whether each child succeeds or fails. The GBG improves the precision with which the teacher conveys and the child receives these rules, and by doing so improves the teacher-child interaction and the child's chances for success. In addition, in GBG trials, the better behaved children were observed to influence and socially integrate the children who behaved less appropriately.

Why the First Grade?

Two considerations recommend the first grade as a setting for preventive interventions:

- Beginning first grade is a major transition for both the child and his or her family;
- First grade is generally the first place where all children—that is, those at all levels of risk of school and behavior problems—can be found. All States in the United States require parents to register their children for first grade with the school district; in many States, this is the first required contact between children and any official system subsequent to birth registration.

The first-grade classroom is well-suited for interventions, such as the GBG, that focus on inculcating the role of students in classrooms. First grade is the first setting outside the home where many children learn the social and behavioral skills they will need to succeed in school. Although some children attend Head Start, kindergarten, or other preschool programs, the length and content of these programs vary.

The first grade is also a particularly appropriate setting in which to provide teachers with tools, as the GBG does, for effective classroom behavioral management. Early in this school year, teachers must organize the classroom, manage children's behavior, and teach rules, but these skills are not intuitive. For example, children in our GBG trial were assigned to first-grade classrooms in a manner that ensured that the classrooms were equivalent with regard to behavior at the start of the school year. However, by the end of the first quarter, when we examined the behavior in the classrooms that had not participated in the GBG, we found that about half were doing relatively well in regard to aggressive and disruptive behavior, while the other half appeared markedly chaotic (Kellam et al., 1998a) (see box).

Social task demands in the classroom include an expectation that children will pay attention, obey rules, learn, and socialize appropriately with their peers and teachers.

TOOLS THAT TEACHERS NEED

School teachers very often report having received little training in tested methods of classroom behavior management. Pre-service teacher training does not emphasize this area, nor does the National Council for Accreditation of Teacher Education (NCATE) require proof of proficiency in this area for schools to be accredited (NCATE, 2008).

Teachers—especially new (National Commission on Teaching and America's Future, 1997) and elementary school teachers (Walter, Gouze, and Lim, 2006)—rate such training as a pressing need. A lack of effective tools to socialize children into the role of student hampers their instruction. The challenge posed by aggressive and disruptive behavior overwhelms many teachers, leading to burnout and resignation from the profession.

The 1985 Baltimore GBG trial provided further evidence that the quality of classroom behavioral management in early grades has far-reaching consequences. Analysis of the data on children who attended standard program (i.e., non-GBG) first-grade classrooms showed a marked influence on the risk of severe aggressive behavior by middle school. Among children in well-managed classrooms, those rated in the top 25 percent for aggressive and disruptive behaviors were up to 2.7 times as likely as the average child to exhibit severe aggressive behavior by middle school. In contrast, in poorly managed classrooms, the risk differential was up to 59 times.

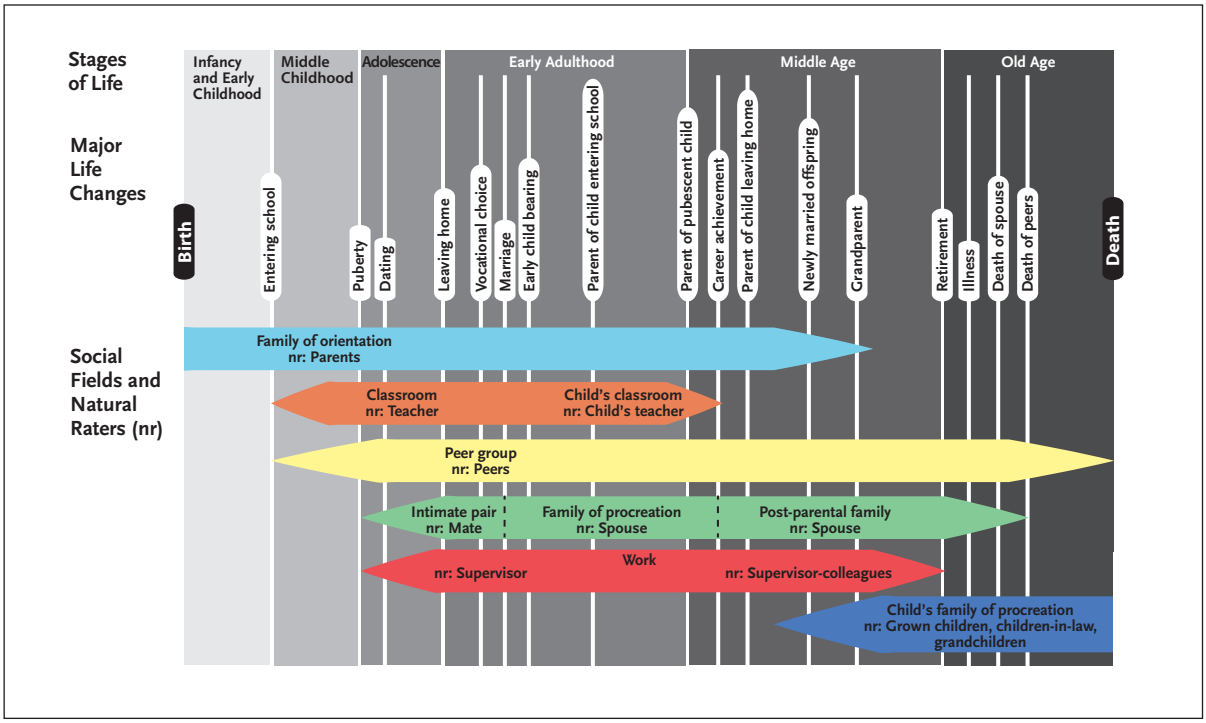
THE THEORY GUIDING THE TRIAL

Prevention trials yield the most insight when they are based on a research-backed theory about causes. For the past 4 decades, life course/social field theory has been a foundation for our research on early developmental risk factors and associated adult problem outcomes and their prevention (Kellam et al., 1976). The theory has pointed to what we needed to measure and what interventions might be effective.

Life course/social field theory provides a dual-faceted view of mental health. In this perspective, adaptation has a social dimension and an individual, psychological dimension.

The social dimension focuses on how an individual is viewed by society, both overall and within specific social contexts. At each stage of life, there are a few main social fields where individuals face social task demands.

FIGURE 1. Life Course/Social Field Concept



For children, the classroom is such a field, where social task demands include an expectation that they will pay attention, obey rules, learn, and socialize appropriately with their peers and teachers. In each social field, the person's ability to meet task demands is assessed or rated by individuals we call natural raters. Teachers and student peers are natural raters in classrooms.

Sometimes this rating process is formal, as in the case of teachers giving grades. At other times, it is informal, as when peers respond to a student. Even when ratings are less formal, however, outcomes such as rejection from the peer group can be very powerful. We call this process of demand and response "social adaptation" and the resulting outcome, "social adaptational status."

An individual may be rated as maladapted for reasons that originate with himself or herself, with the rater, or in the process of demand and response between the two. A first-grader, for example, may behave inappropriately due to a developmental lag in ability to sit still and attend, because the teacher lacks effective methods to socialize students to behave appropriately, or because previous persistent bad behavior has created tension between the teacher and the student.

According to life course/social field theory, improving the way teachers socialize children in the classrooms will result in improved social adaptation of the children in the classroom social field. The theory also predicts that this early improved social adaptation will lead to better adaptation to other social fields over the life course (Figure 1). It is this hypothesis that supports using an intervention like the GBG in first and second grade.

The second dimension in life course/social field theory is the individual's internal condition, or psychological well-being. Depression, anxiety, and thought disorder are examples of poor psychological well-being. Psychological well-being and social adaptational status can reciprocally influence each other over the course of development. For example, receiving poor grades may make a child feel depressed, and depression may make a child more likely to get poor grades. Although the GBG's effects on psychological well-being are beyond the scope of this paper, we have reported on its impact on suicidal thoughts and attempts, and we continue to study this dimension (Kellam et al., 2008; Wilcox et al., 2008).

RESEARCH DESIGN

The trial in the BCPSS tested two classroom interventions. The GBG focused on aggressive and disruptive behavior and is the subject of this paper. An enhanced reading intervention that aimed to improve classroom performance was also tested, but is only mentioned here to provide a complete picture of the study design.

TABLE 1. Young Adult Outcomes in GBG and Standard Classrooms

OUTCOMES	GROUP	GBG CLASSROOM	STANDARD CLASSROOM
Drug abuse and dependence disorders	All males	19 percent	38 percent
	Highly aggressive males	29 percent	83 percent
Regular smoking	All males	6 percent	19 percent
	Highly aggressive males	0 percent	40 percent
Alcohol abuse and dependence disorders	All males and females	13 percent	20 percent
Antisocial personality disorder (ASPD)	Highly aggressive males	40 percent	100 percent
Violent and criminal behavior (and ASPD)	Highly agressive males	34 percent	50 percent
Service use for problems with behavior, emotions, drugs, or alcohol	All males	25 percent	42 percent
Suicidal thoughts	All females	9 percent	19 percent
	All males	11 percent	24 percent

Altogether, 41 classes in 19 schools in five sociodemographically distinct areas of Baltimore participated in the trial. All the students were of low to lower middle socioeconomic status, and 70 percent were African-Americans.

Assignment of Intervention Conditions

Within each urban area, three or four schools were matched and randomly assigned to deliver the GBG, the enhanced reading curriculum program, or no intervention. All students in all schools received the standard first-grade educational program.

Within each intervention school, the principal sequentially assigned students to a first-grade classroom by using an alphabetized list. The research staff then checked and in a few cases adjusted the class rosters with the principal to provide an equivalent distribution of children across classrooms with respect to gender, kindergarten records of behavior, socioeconomic status, and other criteria. Then, within the GBG intervention schools, each first-grade classroom with its teacher was randomly assigned to be a GBG classroom or a standard-program classroom.

This design created three types of control classrooms to compare with the GBG classrooms: (1) standard program classrooms within the schools where the GBG was tested; (2) standard program classrooms within the schools where the enhanced reading curriculum program was tested; and (3) all classrooms within the schools where no intervention was tested. These three controls allowed for extensive analyses that strengthened our confidence in the results. For example, when comparing intervention and standard program classrooms within the GBG schools, we eliminated school and community variation as potential explanations for any differences. Our comparisons of GBG classrooms and standard program classrooms in other schools allowed us to rule out intervention "leakage" into control classrooms within the GBG schools. Using the three kinds of controls, we were also able to collect more information about school- and individual-level variation and compare the consistency of the results across schools and urban areas.

The trial included two consecutive cohorts of children. The first cohort began first grade in 1985. The teachers who had been randomly assigned to deliver the GBG intervention received 40 hours of training in GBG implementation, followed by supportive mentoring and monitoring during the school year. When the students in

the GBG classrooms advanced to second grade the next fall, their new teachers received the same training and support as their first-grade teachers and implemented the intervention again.

Also in the fall of 1986, the second cohort began first grade. The same first-grade teachers who had implemented the GBG in 1985 did so again with their new students. They received little retraining, because we assumed that they would continue implementing the GBG with fidelity.

The resources invested in an intervention can have an effect on outcomes independent of the intervention. To minimize differences due to such effects, we provided all teachers in standard-program classrooms with activities comparable in extent to GBG training and support. The focus of these activities—for example, meetings of teachers from different schools and trips with the children—was not on classroom behavior management.

Teacher ratings have considerable predictive power concerning children's outcomes well into adulthood.

Behavior Measurement

Our primary outcome measures were teachers' ratings of children's social adaptation. A teacher's judgment about how a student is responding to classroom social demands is vitally important, because the teacher strongly influences whether the child continues to the next grade. The teacher is not only a predictor but also a participant in the child's successes or failures. Teacher ratings have considerable predictive power regarding children's outcomes well into adulthood.

The teachers rated the children on the Teacher Observation of Classroom Adaptation–Revised (TOCA–R) scale (Kellam et al., 1976; Werthamer-Larsson, Kellam, and Wheeler, 1991). Ratings were obtained in the fall and spring of grades 1 and 2 and thereafter in the spring of grades 3 through 7.

Each time, a trained interviewer established a relationship of trust with the teacher in a quiet room in the school and then recorded ratings of each child, taking care to spend equal time on each child. Aggressive and disruptive behaviors were specified as: breaks rules, breaks things, fights, harms others, harms property, lies, teases classmates, takes others' property, and yells at others. The teachers' data were validated with other measures, such as classroom peer ratings and observation by independent observers.

Collection of Young Adult Data

When students reached ages 19–21, they were contacted to participate in a 90-minute telephone interview about their social adaptational status within their original family, school, work, intimate relationships, marital family (if any), and peer social fields. They also were asked about any use of services for problems with behavior, emotions, or drugs or alcohol, and about their developmental, behavioral, psychological, and psychiatric status. The Composite International Diagnostic Interview–University of Michigan (CIDI–UM) was used to determine psychiatric diagnoses based on the *Diagnostic and Statistical Manual of Mental Disorders–IV* (DSM–IV) criteria (American Psychiatric Association, 1994; Kessler et al., 1994). Information was also obtained from school and juvenile court and adult incarceration records. A second interview at ages 20–23 was conducted in person to inquire about suicidal thoughts and attempts.

The interviewers did not know which participants had experienced the GBG. Of the students present in the fall of first grade in 1985, 75 percent were interviewed at the young adult followup by telephone or in person. No differences in rates of attrition were found between young adults who were in the GBG classrooms and those in the standard classrooms.

RESULTS

The GBG significantly reduced aggressive and disruptive behavior in primary school classrooms. In the first through sixth grades, students in GBG classrooms, especially the males, exhibited less aggressive and disruptive behavior than those in control classrooms (Dolan et al., 1993). By the spring of sixth grade, males in GBG classrooms who had initially been rated above median levels for aggressive and disruptive behavior had significantly reduced these behaviors (Kellam et al., 1994).

Among females, the levels of aggressive behavior were far lower than for males at the beginning of school and through seventh grade. The intervention did not appear to strongly influence such behavior among females (Kellam et al., 1994; 1998a; 1998b).

Outcomes in Young Adulthood

Male students who had played the GBG in first grade reported significantly fewer problem outcomes at ages 19–21 than their peers who received the standard program. The results were particularly striking for those who had higher levels of aggressive and disruptive behaviors in first grade (Table 1).

Female participants had much lower rates of aggressive and disruptive behaviors in first grade and lower rates

of problem outcomes at ages 19–21. The GBG had little or no statistically significant effect on female outcomes except for suicidal thoughts and, to some extent, alcohol abuse and dependence disorders.

The effectiveness of the GBG was clearest for the most illicit behaviors and disorders—for example, drug abuse and dependence disorders, antisocial personality disorder, and incarceration for violence.

Results for the second cohort, first-graders in 1986, were similar, but there was some reduction of impact. The GBG still appeared to reduce drug abuse and dependence disorders, but instead of the higher risk children benefitting most, the benefit was more general. No significant benefit was seen for alcohol abuse and dependence disorders, regular smoking, or suicidal thoughts or attempts.

RESULTS FROM OTHER GBG TRIALS

Large-scale population-based randomized field trials of the GBG have been completed in three locations and are under way in three others (Mackenzie, Lurye, and Kellam, 2008).

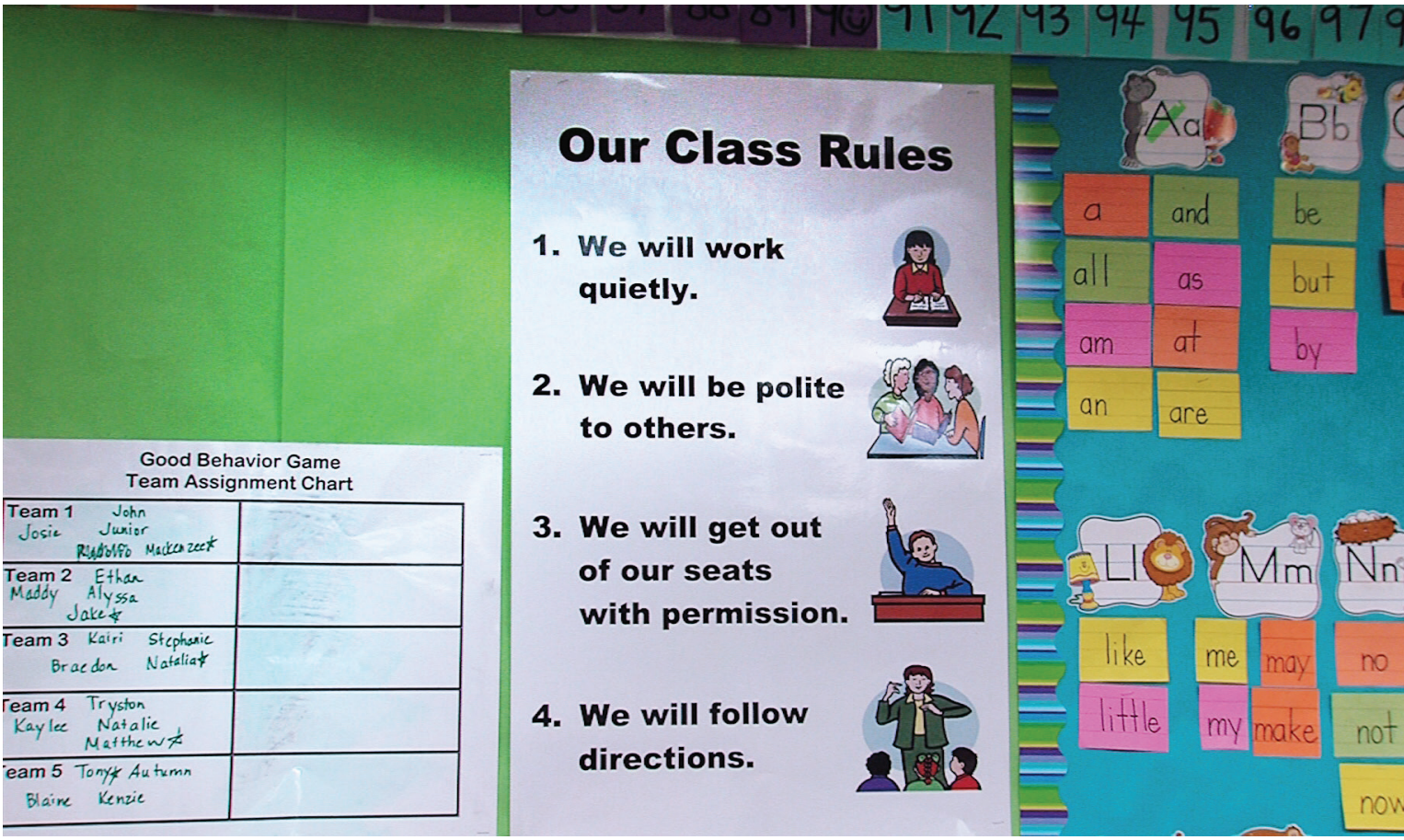
Baltimore, 1990s

A second trial in Baltimore in the early 1990s coupled the GBG with an enhanced curriculum and instruction program. The goal was to improve both behavior and achievement, possibly producing synergism and enhancing and expanding impact. By the end of the first and second grades, the combined intervention had significant positive effects on aggressive and disruptive behavior and achievement (Ialongo et al., 1999). By the end of sixth grade, significant reductions occurred in teacher-rated conduct problems, diagnoses of conduct disorder, school suspensions, use of mental health services, and smoking (Ialongo et al., 2001; Storr et al., 2002; Furr-Holden et al., 2004; Petras, Masyn, and Ialongo, in press).

The GBG significantly reduced aggressive and disruptive behavior in primary school classrooms.

Oregon

The GBG was replicated as a component of a population-based trial designed to target early antecedents of later problem outcomes through a multilevel preventive intervention in the first and fifth grades. The trial, called LIFT (Linking the Interests of Families and Teachers), significantly reduced student aggression during the



intervention period and physical aggression following the intervention (Reid et al., 1999; Stoolmiller, Eddy, and Reid, 2000). Followup analyses 3 years later showed reduced severity of attention deficit disorder behaviors in first-graders and, among fifth-graders, delayed time of first police arrest, association with misbehaving peers, and time to first patterned alcohol and marijuana use (Eddy et al., 2003; 2005; Reid and Eddy, 2002). Further followup of fifth-graders until the end of high school showed significantly reduced overall use of tobacco, alcohol, and illicit drugs (DeGarmo et al., 2009).

Netherlands

The GBG was implemented in the first and second grades in the Netherlands. The results showed that the intervention reduced attention deficit hyperactivity problems. Among the initially more disruptive students, a reduction in conduct problems was seen by the end of third grade (van Lier et al., 2004). By age 10, large reductions were documented in antisocial behavior, and these reductions were associated with lower levels of peer rejection and increased affiliation with nondeviant peers (van Lier, Vuijk, and Crijnen, 2005; Witvliet et al., 2009; van Lier et al., 2011). The GBG also reduced physical and relational victimization at age 10 and major depressive disorder, generalized anxiety disorder, and panic disorder/agoraphobia by age 13 (Vuijk et al., 2007). Further analysis revealed that these reductions in depression and anxiety were mediated by the reductions in relational victimization for girls and physical victimization for boys (Vuijk et al., 2007). Use of tobacco, but not alcohol, between ages 10 and 13 was also reduced among children in GBG classrooms (van Lier, Huizink, and Crijnen, 2009). Later replications of the GBG implemented in the Netherlands showed similar benefits.

Maladapting to the classroom social task demands as early as first grade markedly increases the risk of later serious problems.

Belgium

In an epidemiologically based trial of the GBG in Belgium, Leflot and colleagues reported significant reductions in aggressive and disruptive behavior, increases in on-task behavior, decreases in talking-out behavior, and decreases in the development of oppositional behavior. These results were mediated by decreases in negative teacher remarks (Leflot et al., 2010).

LESSONS LEARNED

The main lesson learned from the GBG trials is that a classroom behavior management intervention directed at aggressive and disruptive behavior in first and second grade can improve children's long-term outcomes. The results of these trials show that such behaviors are malleable to effective universal methods applied with fidelity and consistency.

The improved young adult outcomes of male children who played the GBG point strongly to the conclusion that first-grade classrooms are extremely important to children's development. As many previous studies have reported, maladapting to the classroom social task demands as early as first grade markedly increases the risk of later serious problems. For example, Ensminger and Slusarcick (1992) reported that males' first-grade aggressive behavior coupled with poor academic achievement predicted future school dropout, drug abuse, and criminal behavior. The effect size achieved by the GBG is not surprising when we consider that a child's success or failure in learning to read in the first grade makes a substantial difference to his or her future success in school and beyond.

The impact of the GBG among highly aggressive and disruptive male first-graders—the group most at risk for antisocial and criminal outcomes—adds dramatically to our understanding of such children. The results are consistent with the inference that these behaviors play an etiological role in the development of substance use, antisocial and violent criminal behavior, suicide, and other damaging outcomes.

The minimal impact of the GBG among females calls loudly for further study. Girls' aggressive and disruptive behavior does not appear to have the same importance as boys': It is less prevalent, is less enduring from early to later schooling, and appears less salient for females' long-term development. There is an urgent need for developmental epidemiological studies to understand females' developmental pathways and provide a basis for designing interventions for them.

The Need for Partnerships

Prevention research and programming can succeed only when they are accepted by the community's cultural, social, and political structure (Kellam, 2000). The GBG trials have been possible because their aims have accorded with the mission of the communities in which they were conducted. For example, the BCPSS was willing to commit its resources and expose its students to the research out of deep concern over the problem of socializing young children to be successful students. An equally critical condition for success was that the BCPSS and community exercised oversight over the adaptation of

the GBG for their schools and the design and implementation of the trial. Community oversight can necessitate intense working through of issues, but without it the chances are slim that a prevention program will be adopted, even if it proves effective in trials. In the GBG trial, for example, the families challenged the researchers to show that the randomized design was consistent with the researchers' commitment to carry out the study in accord with the community's values. Ultimately, after intensive discussions and trust building, the families came to see randomization as creating an "even playing field," where every child had the same odds of receiving the GBG or standard program. Moreover, everyone would benefit if the GBG performed as hoped and was accordingly adopted into the curriculum.

This model of partnership for research and later implementation represents the foundation of the next generation of public health, public education, and prevention and treatment research. Researchers will need to understand the mission and vision of local community and institutional leaders, such as ministers and block club leaders, school superintendents, and clinic and other service providers. To ensure that prevention research and programming are conducted and administered with fidelity and continuity over time, researchers will need to integrate "silos," bringing together political and agency leaders at the federal, regional, state, city/county, and local levels. Unfortunately, the formation of such partnerships is still not well-taught in graduate schools.

Networks for Replication

The GBG has now been tested in many pre-post and short-term studies and three large-scale population-based randomized field trials, and further trials are under way in Colorado; Houston, Texas; and Oxfordshire, England. To accelerate these and future replications, and to maximize the information learned from them, we are in the early stage of planning, with NIDA support, a GBG International Network of researchers and their policymaking and institutional partners.

The development of such networks is just beginning in the drug abuse field. However, they are essential for efficiently assessing the effectiveness of prevention interventions through replications on a progressively larger scale and in diverse contexts—to find out what works, for whom, and under what cultural and institutional conditions. Researchers, policymakers, and practitioners will benefit from sharing experiences related to theory, measures, analyses, and obstacles to moving interventions beyond effectiveness trials and into implementation and stages of going to scale. Networks can expedite implementation and expansion into practice by including policymakers and practitioners on the same teams as the researchers.

Integrating Replication and Implementation

Moving the GBG from observational studies to systematic population-based randomized field trials and their long-term outcomes and replication in other sites has taken more than 25 years. For a prevention model developed today, this would be an unacceptably long time. Better theory and new designs and statistical methods make possible more rapid advances from research into practice.

Now is the time to think developmentally and epidemiologically about how an improved health system fits into a broader more functional child development system.

An important new strategy combines replication with expanding previously tested programs system-wide or moving them into new community sites. The first stage of moving a program into new sites or into practice is developing a partnership among community advocates, policymakers, service providers, and the research team that carried out the effectiveness trials. The second stage involves training a cadre of implementers to lead the training of additional implementers. As training proceeds, criteria and instruments used during the effectiveness trial can be streamlined and used to measure the effectiveness of the newly trained implementers. Such a strategy can include the designation of waves of trainees such that some would receive training while others awaited the next wave. Trainees could be randomized if their numbers reached levels that required wait-listing (Brown et al., 2006).

By creating representative stratified samples of schools within a new school district and randomly assigning the trial intervention and control conditions to schools at each stratum, researchers could test training and effectiveness at each stratum in the district. Moving on, the next tier of the stratified sample of schools could be covered in a successive randomized roll-out or "dynamic wait list" design (Brown et al., 2006; 2009). With such designs and methods, the next generation of research, policy, and programming for fostering human development holds great promise.

TOWARD A NEW HUMAN SERVICES SYSTEM

The reform of our health system is at the forefront of our national political and social discourse. Now is the time to think developmentally and epidemiologically, particularly at the community level, about how an improved

health system fits into a broader, more functional child development system. On the basis of our experience with the GBG, we suggest that the potential for such a system depends on expanded school information systems and implementation of staged intervention systems.

The Role of School Information Systems

The GBG trial represents a step toward a long-overdue integration of education research and public health prevention research. Further steps in this direction will be greatly facilitated by expansion of school information systems. As we consider the role of school information systems and community and researcher partnerships, the report entitled *Community-Monitoring Systems: Tracking and Improving the Well-Being of America's Children and Adolescents* (Mrazek, Biglan, and Hawkins, 2004; NIDA, 2007) gives important background information.

The universal strategy is the front line of a system of services that optimizes human development as well as physical health.

Most school information systems primarily monitor academic progress and problems and disciplinary actions. An ideal system would also record each child's progress in emotional and behavioral development, including his or her special needs. The added parameters would inform educators, researchers, and clinicians concerning the child's early risk factors for outcomes such as those measured in the GBG trial as well as family needs and other data. They would support more salient planning for—and responses to—the needs of the individual child, the classroom, and the school.

The No Child Left Behind (NCLB) law presents a unique opportunity to specify both educational and public health needs at the level of demographic epidemiology. NCLB establishes a national, state, and local repository of information that can be analyzed at levels from the national to the community and school district. Depending on the parameters included in NCLB assessments, they can furnish the data for epidemiological studies that show the broad distribution of educational and health-related problems and conditions related to them. These then can be used to plan and implement multilevel community partnerships that include service providers, community advocates, and research teams for testing and implementing effective programs. Communities That Care is one example of a program moving in this direction (Hawkins et al., 2008a; 2008b).

Proper safeguards for confidentiality are possible, as they in fact already exist in a myriad of places where personal data are gathered, such as income taxes, medical records, and mail. Systems of restricted access are needed but should not block the integration of school and social services of other kinds, such as foster care placement, juvenile justice, and child welfare.

The Importance of Staged Interventions

The GBG is a universal intervention; it addresses the entire classroom population, not just those who are at higher risk. In public health, universal programs are usually the strategic first line of defense: Chlorine in drinking water, fluoride in toothpaste, and vaccines against influenza are examples.

Like most universal interventions, the GBG reduced some individuals' risk and averted some adverse outcomes, but not everyone's. In general, children who do not respond well to a universal intervention are candidates for selective prevention (based on persistent risk factors alone), indicated prevention (based on actual symptoms of incipient problems), or treatment.

A coordinated system of staged interventions, consisting of a tested universal intervention backed up by empirically proven group and individual interventions, meets the needs of individuals at all risk levels and stages of problem development. It yields efficiency and economy by differentiating lower risk individuals and higher risk responders from those who need more invasive and costly help. The GBG demonstrated another advantage of universal interventions: It does not single out, and thereby risk stigmatizing, children who manifest aggressive and disruptive behavior. Those who do not respond to universal programs can be reliably identified by their specific needs and enrolled in progressively more selective interventions.

The universal strategy is the front line of a system of services that optimizes human development as well as physical health and is central to the next-stage design of human development services we propose. The most logical place to start building this system is in schools and the agencies that are mandated to serve children with special needs. Pre- and perinatal parental interventions can be important prior prevention services. Once school starts and children become part of the information system, family prevention interventions, developed and tested largely as selective interventions, can be closely integrated as back-up to school-based universal interventions. Partnerships will have to be developed radiating out to community leaders and a broad range of agencies and institutions. The formation of the system and the system itself must be responsive to community values and aspirations, safeguard confidentiality, and ensure proper oversight by appropriate stakeholders.

SUMMARY

The GBG, a universal intervention to manage classroom behavior, reduces schoolchildren's aggressive and disruptive behavior and prevents drug abuse and dependence disorders, violent crime, and other adverse outcomes in young adulthood. Findings from completed and ongoing large-scale GBG trials support the hypothesis that aggressive and disruptive behavior as early as first and second grade plays an etiological role in these adverse outcomes. They also endorse the vision of a national, state, and local human services system, founded in schools, that integrates education and health research and employs a strategy of first-line universal and second-line selective and indicated prevention interventions, backed up by specific treatment programs. The initial work to get this system started has already been done.

ACKNOWLEDGMENTS

During the past 22 years, this research has been supported by National Institute of Mental Health grants R01 MH 42968, P50 MH 38725, R01 MH 40859, and T32 MH018834, with supplements from NIDA for each of the research grants. In addition, work on this paper was supported by grants from NIDA (P30 DA027828 and DA009897). We thank Baltimore City Public School System (BCPSS) staff, especially Ms. Alice Pinderhughes, superintendent during the early years of the study; Dr. Leonard Wheeler, area superintendent; Dr. Patricia Welch, chair of the Board of School Commissioners; and Dr. Carla Ford, head of the kindergarten-through-third-grade reading instruction. We also thank the students who participated and their families. We acknowledge very important contributions of Dr. Jaylan Turkkan in training and supporting the teachers; Dr. Lawrence Dolan as overall intervention chief; Ms. Natalie Keegan, who rallied community support and participation; Dr. James Anthony, who contributed to trial design and early implementation; and Dr. Lisa Ulmer, who led the early assessment team and helped frame the research. We thank the late Dr. Charles R. Schuster, who as director of NIDA augmented the NIMH prevention research center grant with the support needed to carry out the trial. Finally, we are grateful for the editorial contributions by Mr. Matthew Malouf.

CORRESPONDENCE

Sheppard G. Kellam, Professor Emeritus, Bloomberg School of Public Health, The Johns Hopkins University, 624 North Broadway, Room 786, Baltimore, MD 21205; e-mail: skellam@jhsph.edu.

REFERENCES

American Psychiatric Association, 1994. *Diagnostic and Statistical Manual of Mental Disorders* (IV ed.). Washington, DC: American Psychiatric Association.

Barrish, H.; Saunders, M.; and Wolf, M., 1969. Good Behavior Game: Effects of individual contingencies for group consequences on disruptive behavior in a classroom. *Journal of Applied Behavior Analysis* 2(2):119–124.

Brown, C.H., et al., 2006. Dynamic wait-listed designs for randomized trials: New designs for prevention of youth suicide. *Clinical Trials* 3:259–271.

Brown, C.H., et al., 2009. Adaptive designs for randomized trials in public health. *Annual Review of Public Health* 30: 1-25.

DeGarmo, D.S., et al., 2009. Evaluating mediators of the impact of the Linking the Interests of Families and Teachers (LIFT) multimodal preventive intervention on substance use initiation and growth across adolescence. *Prevention Science* 10(3):208–220.

Dolan, L.J., et al., 1993. The short-term impact of two classroom-based preventive interventions on aggressive and shy behaviors and poor achievement. *Journal of Applied Developmental Psychology* 14(3):317–345.

Eddy, J.M., et al., 2003. Outcomes during middle school for an elementary school-based preventive intervention for conduct problems: Follow-up results from a randomized trial. *Behavior Therapy* 34(4):535–552.

Eddy, J.M., et al., 2005. The Linking the Interests of Families and Teachers (LIFT) prevention program for youth antisocial behavior: Description, outcomes, and feasibility in the community. In M.H. Epstein, K. Kutash, A.J. Duchnowski (eds.) *Outcomes for Children and Youth with Emotional and Behavioral Disorders and Their Families: Program and Evaluation Best Practices* (2nd Ed.), Austin, TX: Pro-Ed, Inc., pp. 479–499.

Ensminger, M.E., and Slusarcick, A.L., 1992. Paths to high school graduation or dropout: A longitudinal study of first grade cohort. *Sociology of Education* 65(2):95–113.

Furr-Holden, C.D.M., et al., 2004. Developmentally inspired drug prevention: Middle school outcomes in a school-based randomized prevention trial. *Drug and Alcohol Dependence* 23:149–158.

Hawkins, J.D. et al., 2008a. Early effects of Communities That Care on targeted risks and initiation of delinquent behavior and substance use. *Journal of Adolescent Health* 43(1): 15–22.

Hawkins, J.D. et al.,, 2008b. Effects of social development intervention in childhood 15 years later. *Archives of Pediatric and Adolescent Medicine* 162:1133–1141.

Ialongo, N.S., et al., 1999. Proximal impact of two first-grade preventive interventions on the early risk behaviors for later substance abuse, depression, and antisocial behavior. *American Journal of Community Psychology* 27:599–641.

Ialongo, N., et al., 2001. The distal impact of two first-grade preventive interventions on the conduct problems and disorder in early adolescence. *Journal of Emotional and Behavioral Disorders* 9:146–160.

Kellam, S.G., et al., 1976. *Mental Health and Going to School: The Woodlawn Program of Assessment, Early Intervention, and Evaluation.* Chicago: University of Chicago Press.

Kellam, S.G., et al., 1994. The course and malleability of aggressive behavior from early first grade into middle school: Results of a developmental epidemiologically-based preventive trial. *Journal of Child Psycholology and Psychiatry* 35:259–282.

Kellam, S.G., et al., 1998a. The effect of the level of aggression in the first grade classroom on the course and malleability of aggressive behavior into middle school. *Development and Psychopathology* 10(2):165–185.

Kellam, S.G., et al., 1998b. Effects of improving achievement on aggressive behavior and of improving aggressive behavior on achievement through two preventive interventions: An investigation of causal paths. In B. Dohrenwend (ed.), *Adversity, Stress, and Psychopathology.* London: Oxford University Press, pp. 486–505.

Kellam, S.G., 2000. Community and institutional partnerships for school violence protection. In *Preventing School Violence: Plenary Papers of the 1999 Conference on Criminal Justice Research and Evaluation: Enhancing Policy and Practice through Research,* Vol. 2. Washington, DC: National Institute of Justice, pp. 1–21.

Kellam, S.G., et al., 2008. Effects of a universal classroom behavior management program in first and second grades on young adult behavioral, psychiatric, and social outcomes. *Drug and Alcohol Dependence* 95(Suppl 1):S5–S28.

Kessler, R., et al., 1994. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States: Results from the National Co-morbidity Survey. *Archives of General Psychiatry* 51(1):8–19.

Leflot, G., et al., 2010. The role of teacher behavior management in the development of disruptive behaviors: An intervention study with the Good Behavior Game. *Journal of Abnormal Child Psychology* 38(6):869-882.

Mackenzie, A.C.; Lurye, I.; and Kellam, S.G., 2008. History and evolution of the Good Behavior Game. Supplementary material for: Kellam, S.G., et al., 2008. Effects of a universal classroom behavior management program in first and second grades on young adult behavioral, psychiatric, and social outcomes. *Drug and Alcohol Dependence* 95(Suppl 1):S5–S28.

Mrazek, P.; Biglan, A.; and Hawkins, J.D., 2004. *Community-Monitoring Systems: Tracking and Improving the Well-Being of America's Children and Adolescents.* Society for Prevention Research, available at *www.preventionresearch.org/CMSbook.pdf.*

National Commission on Teaching and America's Future, 1997. *Doing What Matters Most: Investing in Quality Teaching.* New York: National Commission on Teaching and America's Future.

National Council for Accreditation of Teacher Education, 2008. *Professional Standards for the Accreditation of Teacher Preparation Institutions.* Available at *www.ncate.org/documents/standards/NCATE%20Standards%202008.pdf.*

National Institute on Drug Abuse, 2007. *Community Monitoring Systems: Tracking and Improving the Well-Being of America's Children and Adolescents.* National Institutes of Health, U.S. Department of Health and Human Services, available at *www.nida.nih.gov/pdf/cms.pdf*

Petras, H.; Masyn, K.; and Ialongo, N., in press. The developmental impact of two first grade preventive interventions on aggressive/disruptive behavior in childhood and adolescence: An application of latent transition growth mixture modeling. *Prevention Science.*

Reid, J.B., and Eddy, J.M., 2002. Preventative efforts during the elementary school years: Linking the Interests of Families and Teachers (LIFT) project. In J.B. Reid, G.R. Patterson, J. Snyder (eds.), *Antisocial Behavior in Children and Adolescents: A Development Analysis and Model for Intervention.* Washington, DC: American Psychological Association, pp. 219–233.

Reid, J.B., et al., 1999. Description and immediate impacts of a preventive intervention for conduct problems. *American Journal of Community Psychology* 27:483–517.

Stoolmiller, M.; Eddy, J.M.; and Reid, J.B., 2000. Detecting and describing preventive intervention effects in a universal school-based randomized trial targeting delinquent and violent behavior. *Journal of Consulting and Clinical Psychology* 68:296–306.

Storr, C.L., et al., 2002. A randomized controlled trial for two primary school intervention strategies to prevent early onset tobacco smoking. *Drug and Alcohol Dependence* 66:51–60.

van Lier, P.A.C., et al., 2004. Preventing disruptive behavior in elementary schoolchildren: Impact of a universal classroom-based intervention. *Journal of Consulting and Clinical Psychology* 72:467–478.

van Lier, P. A. C.; Huizink, A.; and Crijnen, A., 2009. Impact of a preventive intervention targeting childhood disruptive behavior problems on tobacco and alcohol initiation from age 10 to 13 years. *Drug and Alcohol Dependence* 100(3):228–233.

van Lier, P.A.C.; Vuijk, P.; and Crijnen, A.A.M., 2005. Understanding mechanisms of change in the development of antisocial behavior: The impact of a universal intervention. *Journal of Abnormal Child Psychology* 33:521–535.

Vuijk, P., et al., 2007. Testing sex-specific pathways from peer victimization to anxiety and depression in early adolscents through a randomized intervention trial. *Journal of Affective Disorders* 100:221–226.

Walter, H.J.; Gouze, K.; and Lim, K.G., 2006. Teachers' beliefs about mental health needs in inner city elementary schools. *Journal of the American Academy of Child and Adolescent Psychiatry* 45:61–68.

Werthamer-Larsson L.; Kellam, S.G.; and Wheeler, L., 1991. Effect of first grade classroom environment on child shy behavior, aggressive behavior, and concentration problems. *American Journal of Community Psychology* 19(4):585–602.

Wilcox, H.C., et al., 2008. The impact of two universal randomized first- and second-grade classroom interventions on young adult suicide ideation and attempts. *Drug and Alcohol Dependence* 95(Suppl 1):S60–S73.

Witvliet, M., et al., 2009. Testing links between childhood positive peer relations and externalizing outcomes through a randomized controlled intervention study. *Journal of Consulting and Clinical Psychology* 77(5):905–915.