

Abstract Title Page
Not included in page count.

Title: Measurement of Child Behavior via Classroom Observations in the Good Behavior Game Professional Development Models Randomized Control Trial

Authors and Affiliations:

Anja Kurki, Ph.D., American Institutes for Research, email: akurki@air.org

Wei Wang, Ph.D. University of South Florida, email: wwang@health.usf.edu

Yibing Li, Ph.D., American Institutes for Research: email: yli@air.org

Jeanne Poduska, Sc.D, American Institutes for Research, email: jpoduska@air.org

Acknowledgement:

This work is jointly supported by IES grant R305A090446 (PI Poduska), NIDA grant R01 DA030452 (PI Poduska).

Abstract Body

Limit 4 pages single-spaced.

Background / Context:

Description of prior research and its intellectual context.

The Good Behavior Game (GBG) is a classroom-based behavior management strategy aimed at reducing aggressive/disruptive behavior and socializing children into the role of student. GBG, delivered in first and second grades, has been shown to reduce rates of substance abuse and other deleterious outcomes into young adulthood (Brown, C.H. et al 2007, Ialongo, et al. 2001, Kellam, S.G. et al, 2008 and Poduska, J.M. et al. 2008, Kellam, S.G. et al 1998, Ialongo, N.S. et al. 1999), especially for students that were rated as aggressive, disruptive at entry into first grade/baseline by their teachers. Despite GBG being recognized as an evidence-based prevention program (National Registry of Evidence-based programs and Practices, NREPP, <http://nrepp.samhsa.gov/>), published data to date is mostly limited to African American and Anglo populations that are native English speakers.

This study is a four-year IES Goal 3 project (R305A090446, PI Poduska) that was designed to understand the level of professional development (PD) required for teachers to implement and sustain the Good Behavior Game (GBG), a universal preventive intervention. Two models of PD are being compared with each other and with a control condition, in a randomized field trial. **GBG Basic**, which provides group-based pre-implementation training to teachers supplemented by a group-based booster session, and **GBG w Coach**, which has the same activities under **GBG Basic** plus a coach who works directly in the classroom with the teacher, This trial takes place in multilingual context of Houston Independent School district.

Purpose / Objective / Research Question / Focus of Study:

Description of the focus of the research.

Aggressive and disruptive behaviors, as early as first grade, are confirmed antecedents for social and behavioral problems later in life (Kellam et al., 2008). Accurate and meaningful measurement of child behavior is essential in determining the effectiveness of preventive interventions. Previous trials testing the effectiveness of GBG alone or in combination with other components (Ialongo et al., 1999) used a focal child approach for classroom observations, in which each child in a class was observed for a short, discrete time interval. At a given time point, only one child was observed; during an entire observation, each child was observed for 2-6 minutes depending on the size of the classroom and the length of the observation. This approach may fail to produce an adequate description of child behaviors since only the behavior of the focal child will be recorded at the particular time point. To accurately and meaningfully capture behaviors at classroom level, we developed a new classroom observation approach that uses time sampling to continuously observe all children minute by minute.

The purpose of the paper is twofold. First, we will describe the observation protocol and its reliability. The developed protocol documents student behavior (off-task, disruptive, aggressive and social isolation), classroom context (instructional format, subject area), overall level of classroom behavior (a classroom scan every 5 minutes) and teachers' instructional and classroom management style (punitive, neutral, positive). Second, we will describe preliminary child behavior results based on classroom observations conducted for the aforementioned GBG trial. Specifically, regarding the protocol itself, we will report on inter-rater reliability of

observers, average lengths of observations, average number of students in the classroom, the typical instructional subject area and format. Regarding student behavior we will report the rates of off-task and disruptive and aggressive behavior. In addition to descriptive analyses we will conduct analysis to see whether the observation data correlates with our other student-level outcome based on teacher interviews. We will also conduct regression analysis to see whether student or classroom/teacher characteristics are related to observed student behaviors.

Setting:

Description of the research location.

The study is taking place in partnership with the Houston Independent School District (HISD) and the Houston Federation of Teachers—the teachers’ union. In 2010-2011 65 percent of elementary school enrolled students were Hispanic, 24 percent were African American and 7 percent were White. In addition, 43 percent of students were classified as Limited English Proficiency (LEP) and 85 percent were classified as having free/reduced lunch status. The study sample includes schools with more Hispanic and fewer White students, compared to the district average. In addition, the sample schools appear to have higher percentage of free/reduce lunch students, LEP students, bilingual students and students At-Risk than an average elementary school in the district which is in keeping with the sampling strategy for the trial; we targeted high need schools using the district’s At-Risk designation, measures of academic achievement, and free/reduced lunch status.

Population / Participants / Subjects:

Description of the participants in the study: who, how many, key features, or characteristics.

The study will include three cohorts of students; this paper will be based on the data collected from the first cohort of teachers and students that participated in the study during the school years 2010-2011 and 2011-2012. This cohort includes 18 schools, 71 classrooms and 1401 students. Table 1 shows the number of bilingual and regular education classrooms and students included in the first cohort.

Intervention / Program / Practice:

Description of the intervention, program, or practice, including details of administration and duration.

The Good Behavior Game is a group-based classroom behavior management strategy that promotes classmate/peer concern for each child’s positive behavior by rewarding teams with appropriate behavior. GBG also provides teachers with additional skills in classroom behavior management resulting in fewer negative interactions, reductions in aggressive, disruptive behavior, and an increase in time spent on-task. As tested in this trial GBG is built around four core elements integrating (1) classroom rules, (2) team membership, (3) monitoring of behavior, and (4) positive reinforcement to individuals and the group. GBG, delivered in first and second grades, has been shown to reduce rates of substance abuse and other deleterious outcomes into young adulthood (Kellam, S.G. et al, 2008 and Poduska, J.M. et al. 2008, Petras, H. et al 2008). However, although GBG has been proven as an effective intervention, no rigorous studies have been conducted before about the necessary supports teachers would need for high level of implementation or implementation over time. In this trial two different training models are tested: the *GBG Basic* option (2 days of initial training and one day of booster training) and the

GBG with Coach (2 days of initial training and one day of booster training and 90 minutes of coaching every two weeks for one school year).

An added complexity for this trial is the multilingual environment in which the trial takes place. The group based training was provided by English; teachers in bilingual classrooms who were assigned to the *GBG with Coach* condition were provided a bilingual coach. All classroom materials were provided both in English and Spanish in bilingual classrooms.

Research Design:

Description of the research design.

The study is a randomized control trial in which students are first randomly assigned to classrooms and classrooms are randomly assigned to two treatment conditions (*GBG basic*, *GBG with Coach*) or a control condition. The original plan was to recruit schools that have three first grade classrooms to create a balanced design. However, due to the large number of bilingual classrooms and the need to compare the intervention's effect by comparing randomly assigned classrooms of the same type (bilingual v. regular education), the study team adopted an unbalanced design. We created "virtual schools" by grouping classrooms within schools which had the same language of instruction; random assignment of children and teachers was conducted within the virtual schools.

Data Collection and Analysis:

Description of the methods for collecting and analyzing data.

The study engaged in comprehensive data collection both regarding implementation of the GBG and student outcomes. Table 2 presents data collected for this study. In this paper we will focus on classroom observation data that was collected during the implementation year, the year when teacher are trained to use GBG and, if randomly assigned to *GBG with Coach*, receive coaching support. We will also use teacher interview data about student behavior, which were also collected each fall and spring (TOCA Interview Protocol).

We will analyze data that were collected in fall (baseline: 2010 and 2011) and in spring (outcome: 2011 and 2012). We believe that the systematic variation in student behavior exists. Therefore we will present descriptive data for the overall sample as well as by classroom and by selected teacher characteristics, such a bilingual vs. general education classrooms and high or low levels of teacher self-reported classroom management efficacy (Ohio State Teacher Sense of Self-Efficacy (OSTES) or professional burnout (The Maslach Burnout Inventory (MBI). We will also present descriptive statistics by gender for the students. To explore whether observed student behaviors are associated with teachers' reports about students' behavior, we will conduct correlational analyses.

Finally, we will present preliminary impact results for the first cohort of students: whether classrooms' intervention condition is related to observed student behavior (aggressive/disruptive behavior, off-task behavior). We will also conduct regression analysis to explore what student and classroom/teacher level factors predict observed student behavior. The regression analyses will use HLM approach that acknowledges the clustered structure of the data (students within classrooms, classrooms within schools).

Findings / Results:

Description of the main findings with specific details.

Preliminary analyses regarding inter-rater reliability suggest that the observers reached 80% or higher agreement levels for student behavior (percentage agreement) and other information collected during the observations. The students were observed in 5 minute segments and on average 43.8 (std = 9.1) and 41.1 (std = 9.4) minutes during the fall in the morning and afternoon, respectively.

Analysis regarding student behavior for baseline shows that 51.0 and 52.3 percent of students were observed as off-task, and 17.6 and 24.9 percent as disruptive during morning and afternoon observations. As would be expected, aggressive behavior was observed rarely, 2.2 and 3.0 percent of students were identified as engaged in aggressive behaviors during morning and afternoon observations (Table 3). The levels of off-task, disruptive, and aggressive behaviors decrease over the school year as would be expected due to maturation of student. Table 3 shows the levels of these behaviors for the spring observations.

Although instances of aggressive behavior are low, there is a positive correlation between teacher ratings of aggressive/disruptive behavior and the observed aggressive behavior. Similarly, there is a positive correlation between the number of instances students are off-task and teachers' ratings about hyperactiveness/impulsivity and attention/concentration (Tables 4a and 4b).

The impact analysis will be conducted during the summer 2013 when we have finalized student demographic student data from the HISD.

Conclusions:

Description of conclusions, recommendations, and limitations based on findings.

Aggressive and disruptive behaviors, as early as first grade, are confirmed antecedents for social and behavioral problems later in life (Kellam et al., 2008). Accurate and meaningful measurement of child behavior is essential in determining the effectiveness of preventive interventions. In this paper we describe a protocol that was developed for the GBG intervention to capture student behavior in classrooms and present preliminary results for the first cohort of students participating in the GBG PD Models trial. In addition to documenting off-task, disruptive, aggressive and isolated behaviors, the observation protocol captures classroom level behavior management in general and additional instruction related factors such as subject area and instructional format. The inter-rater reliability (measured as percentage of agreement) is 80% or higher for student related behaviors and other classroom management related information (classroom scan, classroom level items about classroom atmosphere and quality). The student behavior outcomes collected through classroom observations also positively correlate with teachers' reports about student behavior. The impact analysis will be conducted during summer 2013 when student demographic data will be available from the HISD.

Based on the preliminary results the protocol is able to capture student behaviors with adequate reliability and the protocol provides a wealth of information about classrooms that can be used to describe student and teacher behaviors as well as the instructional context of the classrooms.

Appendices

Not included in page count.

Appendix A. References

References are to be in APA version 6 format.

- Brown, C.H., Kellam, S.G., Ialongo, N., Poduska, J., & Ford, C. (2007). Prevention of aggressive behavior through middle school using a first-grade classroom-based intervention. In M. T. Tsuang, W.S. Stone, & M. J. Lyons (eds.), *Recognition and Prevention of Major Mental and Substance Use Disorders* (pp 347-369). Washington DC: American Psychiatric Publishing, Inc.
- Ialongo, N., Poduska, J., Wethamer, L., Kellam, S. (2001). The distal impact of two first-grade preventive interventions on conduct problems and disorder in early adolescence. *Journal Of Emotional And Behavioral Disorders*, 9, 146-160.
- Ialongo, N. S., Werthamer, L., Kellam, S., Brown, C. H., Wang, S., & Lin, Y. (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.
- Kellam, S.G., Brown, C.H., Poduska, J.M., Ialongo, N.S., Wang, W., P. Toyinbo, Petras, H., Ford, C., Windham, A., Wilcox, H.C. (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), 1-28.
- Kellam, S. G., Ling, X., Merisca, R., Brown, C. H., & Ialongo, N. (1998). 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, 165-185. See also the erratum: Kellam, S. G., Ling, X., Merisca, R., Brown, C. H., & Ialongo, N. (2000). 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. *Development and Psychopathology*, 12
- Maslach C, Jackson SE, Leiter MP. Maslach burnout inventory: Third edition. In: Wood RJ, editor. *Evaluating stress: A book of resources*. Lanham, MD US: Scarecrow Education; 1997. p. 191-218.

- Petras, H., Kellam, S.G., Brown, C.H., Muthen, B.O., Ialongo, N.S., Poduska, J.M. (2008). Developmental epidemiological courses leading to antisocial personality disorder and violent and criminal behavior: Effects by young adulthood of a universal preventive intervention in first- and second-grade classrooms. *Drug and Alcohol Dependence*, 95(Suppl. 1), 45-59.
- Poduska, J.M., Kellam, S.G., Wang, W., Brown, C.H., Ialongo, N.S., Toyinbo, P. (2008). 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 and Alcohol Dependence*, 95(Suppl. 1), 29-44.
- Tschannen-Moran M, Woolfolk Hoy A, Hoy WK. Teacher efficacy: Its meaning and measure. *Review of Educational Research*. 1998; 68(2): 202-248.

Appendix B. Tables and Figures

Not included in page count.

Table 1. Number of Classrooms and Students in Cohort 1 by Intervention and Language Status

	GBG w Coach	GBG Basic	Standard	Total
Bilingual Classrooms				
Number of Teachers	10	14	12	36
Number of Students	206	297	235	738
Regular Education Classrooms				
Number of Teachers	14	13	8	35
Number of Students	262	259	142	663
Total	468	556	377	1401

Table 2. Data Collection Framework in the GBG professional Development Models Trial

Construct	Data Collected	Timing
Student Outcomes: aggressive, disruptive and socially isolated behavior	Teacher Interview: Teacher Observation of Classroom Adaptation (TOCA)	Fall 2010/11 (baseline), Spring 2011/12 (Outcome)
	Classroom Observations: minute-by-minute documentation of off-task, aggressive and socially isolated behavior	Fall 2010/11 (baseline), Spring 2011/12 (Outcome)
	Extant Data: Student Achievement	Fall 2011, 2012, 2013
Teacher Practices: teachers positive/negative instructions regarding instruction/behavior	Classroom Observations	Fall 2010/11 (baseline), Winter 2011/12 (Outcome), Spring 2011/12 (Outcome)
Implementation of GBG and Coaching: training dosage, dosage and quality coaching process, dosage and quality of GBG implementation in classrooms	Coaching process data: Fidelity Checklist, Professional Development Plan, Progress Reports, Probe Reports	Fall 2010/11 – Spring 2011/12 – web-based data collection
	Classroom observations in GBG classrooms by coaches: fidelity of implementation	Winter 2011/12, Spring 2011/12
	Classroom observations by independent observers	Fall 2010/11, Winter 2011/12, Spring 2011/12

	regarding GBG implementation	
	Teacher reports of use of GBG in classrooms: implementation dosage	Fall 2010/11 – Spring 2011/12
	Qualitative interviews about GBG Adaptation	November-December 2011/12 April-May 2012/13
Multi-level Contextual Factors <ul style="list-style-type: none"> • Program • Teacher • Principal/School 	Teacher surveys and principal surveys	Fall 2010/11, Winter 2011/12 (only for teachers implementing GBG), Spring 2011/12

Table 3. Description of Student Behavior in Classrooms, Cohort 1

Student Behavior	Fall Observations		Spring Observations	
	A.M.	P.M.	A.M.	P.M.
Off-Task	51.0%	52.3%	36.9%	44.17%
Disruptive	17.6%	24.9%	10.4%	14.8%
Aggressive (Verbal or Physical)	2.2%	3.0%	0.79%	2.7%

Table 4a. Correlations Between Observed and Teacher Reported Aggressive Student Behavior, Fall Baseline Data for Cohort 1, Control Students

	PM Aggressive Behavior	AM Aggressive Behavior	Aggressive/Disruptive Behavior (Teacher Report)
PM Aggressive Behavior	1	0.00118	0.11752
		p = 0.9688	p = <.0001
	(N=1210)	(N=1110)	(N=1184)
AM Aggressive Behavior	0.00118	1	0.11688
	p = 0.9688		p = <.0001
	(N=1110)	(N=1189)	(N=1162)
Aggressive/Disruptive Behavior (Teacher Report)	0.11752	0.11688	1
	p = <.0001	p = 0.9688	
	(N=1184)	(N=1162)	(N=1302)

Table 4b. Correlations Between Observed Off-Task and Teacher Reported Off-Task Types Student Behavior, Fall Baseline Data for Cohort 1, Control Students

	AM Off-Task	PM Off-Task	Attention/ Concentration	Hyperactiveness /Impulsivity
AM Off-Task	1	0.21279	0.2011	0.25197
		p = <.0001	p = <.0001	p = <.0001
	(N=1189)	(N=1110)	(N=1162)	(N=1162)
PM Off-Task	0.21279	1	0.17051	0.14064
	p <.0001		p = <.0001	p = <.0001
	1110	1210	1184	1184
Attention/Concentration	0.2011	0.17051	1	0.57924
	p <.0001	p <.0001		p <.0001
	(N=1162)	(N=1184)	(N=1302)	(N=1302)
Hyperactiveness/ Impulsivity	0.25197	0.14064	0.57924	1
	p <.0001	p <.0001	p <.0001	
	(N=1162)	(N=1184)	(N=1302)	(N=1302)