

Getting Teachers in on the Act: Evaluation of a Theater- and Classroom-Based Youth Violence Prevention Program

MARLA ZUCKER, JOSEPH SPINAZZOLA, AMIE ALLEY POLLACK,
LAUREN PEPE, and STEPHANIE BARRY

Trauma Center at Justice Resource Institute, Brookline, Massachusetts, USA

LYNDA ZHANG

Harvard Business School, Harvard University, Boston, Massachusetts, USA

BESSEL VAN DER KOLK

Trauma Center at Justice Resource Institute, Brookline, Massachusetts, USA

This study replicated and extended our previous evaluation of Urban Improv (UI), a theater-based youth violence prevention (YVP) program developed for urban youth. It assessed the replicability of positive program impacts when implemented by nonprogram originators, as well as the utility of a comprehensive version of the UI program that included a classroom-based teacher curriculum (TC). Outcomes compared 260 elementary school students in classrooms that received basic UI, the TC, or the comprehensive program (UI + TC), or to serve as matched control classrooms. Findings revealed that UI can be generalized to new troupes, the comprehensive program demonstrates an additive effect over UI alone, and the TC shows some promise as an easily disseminated, stand-alone YVP program.

KEYWORDS youth violence, prevention, theater, elementary school, outcome evaluation, replication

Received February 27, 2009; accepted July 30, 2009.

The research presented in this paper was funded in part by grants from the Hamilton Fish Youth Violence Prevention Consortium administered through the Office of Juvenile Justice and Delinquency Prevention and from the Roy A. Hunt Foundation.

Address correspondence to Marla Zucker, PhD, Trauma Center at Justice Resource Institute, 1269 Beacon St., Brookline, MA 02446, USA. E-mail: mzucker@jri.org

Violence carried out and/or witnessed by adolescents is a public health problem that has been linked to not only increases in physical injury, but also problems in social, cognitive, and behavioral functioning (Botvin, Griffin, & Nichols, 2006; Margolin & Gordis, 2004). Over the past two decades, the school setting has become a critical nexus for both the incidence of violence as well as for the implementation of youth violence prevention programs (Eisenbraun, 2007). In response to the alarming rise in school and community violence, the federal government instituted the Safe Schools/Healthy Students Initiative (SS/HS) to provide grants to develop real-world knowledge about what works best to reduce school violence and to support a comprehensive approach to prevent youth violence. While such programming has most often been designed to decrease youth aggression and increase prosocial behaviors, innovative research sponsored through this initiative has also addressed the impact of youth violence and preventive programming on less-explored domains such as scholastic engagement, leadership, and mental health (Twemlow, Fonagy, & Sacco, 2001).

Thanks in large part to SS/HS and other government initiatives, a body of research has significantly advanced the evidence base regarding youth violence prevention (YVP) programming. Studies have examined the scope of violence prevention initiatives, including the advantages of selective versus universal/whole-school initiatives (e.g., Farrell, 2008), as well as evaluation of the quality of program implementation (e.g., Hussey & Flannery, 2007). Researchers have also examined the impact of YVP programs on a range of outcomes, from academic achievement (e.g., Chen, 2007) to mental health outcomes (e.g., Stein et al., 2003; Zyromski, 2007) to engagement in violent and aggressive acts (e.g., Hussey & Flannery, 2007), using both qualitative and quantitative data (e.g., Edwards, Hunt, Meyers, Grogg, & Jarrett, 2005). Moreover, a range of potential risk and protective factors to youth violence have been addressed (e.g., Cross, Mohajeri-Nelson, & Newman-Gonchar, 2007). Study design has also varied widely, from naturalistic and quasi-experimental studies (e.g., Massey, Boroughs, & Armstrong, 2007) to randomized, controlled trials (e.g., Stein et al., 2003), and concordance among different sources of outcome data, such as parent versus student ratings, has been explored (e.g., Holt, Kantor, & Finkelhor 2009). Ethnocultural considerations in YVP has been studied (e.g., Zyromski, 2007), as well as gender differences in response to YVP programming (e.g., Griffin, Chen, Eubanks, Brantley, & Willis, 2007). Finally, gains have been made in achieving and evaluating school-community partnerships that work collaboratively to reduce youth violence (e.g., Massey et al., 2007).

Research indicates that school-based prevention programs should intervene as early as possible and that those targeting primary prevention aims (i.e., to keep problems from emerging), should be universal, with all students participating in their routine classroom settings (Sprague & Walker,

2000). Such programs are postulated to benefit overall school functioning, as well as reduce delinquency and antisocial behavior (Hahn, Crosby, Moscicki, Stone, & Dahlberg, 2007). Guidelines for effective YVP programming recommended by Greenberg et al. (2003), Nation et al. (2003), and Weissberg, Kumpfer, and Seligman (2003) included: (a) providing multiple components and targeting multiple systems across multiple years; (b) providing opportunities for active practice of learned skills; (c) fostering positive, supportive relationships between students and school staff; (d) being developmentally and culturally appropriate; (e) beginning prior to the development of problem behaviors; (f) being implemented by well-trained and well-supported staff; and (g) incorporating outcome evaluation and continuous monitoring of effectiveness. Theater-based strategies for YVP have been posed as especially promising (Office of Juvenile Justice and Delinquency Prevention, 1995), as many incorporate these best practice recommendations and may be particularly effective for engagement and learning among urban youth who are regularly faced with violent scenarios.

In a previous study (Kisiel et al., 2006), we conducted an independent evaluation of Urban Improv (UI), a community-based model of a school- and theater-based YVP program that has been in operation in the Boston Public Schools since 1993. UI utilizes structured theater improvisation to teach decision-making, conflict resolution, problem solving, and impulse control skills. This evaluation assessed the effectiveness of the UI program for at-risk fourth graders by examining its impact on three outcome domains related to violence exposure, perpetration, and associated youth well-being and functioning: (a) aggressive and externalizing behaviors; (b) prosocial behaviors, including cooperation, assertiveness, and self-control; and (c) scholastic attention and engagement. In this study, the program's originators implemented UI. Results of this quasi-experimental, controlled outcome study using teacher reports showed that it prevented the increase or new onset of aggressive behaviors and increased prosocial behaviors and classroom engagement. In contrast, comparison students demonstrated increases in teacher-reported aggression, hyperactivity, and withdrawal, and deterioration of prosocial behaviors.

This basic UI program met most of the best practice guidelines for YVP, but did not provide opportunities for practice outside of sessions or directly support the cultivation of stronger relationships between students and teachers, as teachers are silent observers of the standard UI program, which occurs in the community, outside the classroom setting. To enhance the UI program as a YVP initiative, therefore, we developed a classroom-based, teacher-led component to extend the application of UI violence or problem resolution strategies within a broader context and to cement learning through increased practice. *Urban Improv for the Classroom: 4th Grade Teacher Curriculum* (referred to as TC in this article; Zucker, Villaflor, Holden, Jones, & Spinazzola, 2006) was developed by a team of fourth-grade

Boston public school teachers, UI actor/educators, and trauma center psychologists. This adaptation of the UI program has the additional advantage of allowing for more efficient dissemination, as it requires less extensive staff training and fewer resources by drawing from UI's core components and building them into the natural setting of the classroom.

The present study replicated and extended our previous evaluation of UI (Kisiel et al., 2006). It also builds directly from the broad body of research that now exists on YVP, as the goal of the intervention was to enhance community-school partnerships in the prevention effort, attending to a range of outcomes, both prosocial and antisocial, the concordance between sources of outcome data (student vs. teacher-report), and the consistency with which the model was implemented across classrooms. The general aim of the present study was to assess the transferability of the UI program to new staff and school districts by examining whether it continues to demonstrate positive impacts when implemented by nonprogram originators. The present study evaluated the comprehensive version of the UI program, which included both the basic UI protocol and the UI TC. Specifically, we assessed: (a) whether the comprehensive UI program has an additive effect on the three outcome domains (aggression, prosocial behaviors, and scholastic attention and engagement) compared to basic UI; and (b) whether the UI TC demonstrates utility as a stand-alone YVP program when the comprehensive program is not administered.

METHOD

Participants

Participants were fourth-grade students drawn from six inner-city schools within one school district. Intervention schools were chosen based on prior collaborations and at-risk demographic factors primarily related to school location (e.g., urban, high-crime location of school system). Control classrooms were drawn from the same schools as intervention classrooms. Inclusion criteria for the study were: (a) enrollment in the fourth grade at a Boston Public School, (b) enrollment in a UI and/or TC class or a selected control classroom, (c) consent from legal guardian, (d) student agreement to participate and ability to provide assent, and (e) English language proficiency. Recruitment was sought from all students in a given class.

Participants were 260 students from 15 fourth-grade classrooms: four classes ($n = 71$) received the comprehensive intervention (basic UI and the TC); four classes ($n = 70$) participated in basic UI; three classes ($n = 43$) received the TC as a stand-alone intervention; and four classes ($n = 76$) served as controls. Participants ranged in age 9–11 years old ($M = 9.52$, $SD = 0.63$). Gender was approximately evenly distributed (55.8% female and 44.2% male). Ethnic/racial distribution was predominantly ethnic minority:

TABLE 1 Descriptive Variables by Intervention Group

Demographic variable	Comprehensive UI (Basic + TC)	Basic UI	UI TC	Control
Total <i>n</i>	71	70	43	76
Mean age (<i>SD</i>)	9.3 (.53)	9.7 (.67)	9.4 (.58)	9.6 (.66)
Gender (%)				
Female	53.5	57.1	53.5	57.9
Male	46.5	42.9	46.5	42.1
Ethnicity (%)				
African-American	52.2	32.9	42.9	18.3
Caucasian	7.2	5.7	0.0	2.8
Hispanic/Latino	15.9	42.9	28.6	62.0
Asian	0.0	0.0	2.4	1.4
Native American	1.4	4.3	0.0	1.4
Bi/Multiracial	7.0	8.6	18.6	11.3
Other	15.9	5.7	7.2	2.8

Note. UI = Urban Improv; TC = teacher curriculum.

Hispanic/Latino (38.5%), African American (35.7%), Bi/multiracial (10.7%), Caucasian (4.4%), Native American (2.0%), Asian (0.8%), and Other (7.9%). Demographic variables by intervention group are reported in Table 1.

Measures

SOCIAL SKILLS RATING SYSTEM: ELEMENTARY LEVEL (SSRS)

The SSRS (Gresham & Elliot, 1990) was used to measure participants' aggression or externalizing behaviors, social skills, and attention or engagement. Both the student (SSRS-S) and teacher (SSRS-T) report versions were used. The SSRS-S is a 34-item self-report measure scored on a 3-point response scale. Items load onto four subscales: Cooperation, Assertion, Empathy, and Self-Control. The SSRS-T consists of 30 items related to social skills (cooperation, assertion, self-control) and 12 items representing problem behaviors (internalizing, externalizing, hyperactivity). For the purposes of this study, the SSRS-T Externalizing scale was used as an index of aggression or disruptive behaviors, and the SSRS-T Internalizing and Hyperactivity scales were used as indices of classroom attention and engagement.

The SSRS-T has demonstrated good internal consistency, with coefficient alphas of .93 (females) and .94 (males) for the Social Skills scale, .87 (females) and .88 (males) for the Problem Behaviors scale, and .96 (females) and .95 (males) for the Academic Competence scale (Gresham & Elliot, 1990). The internal consistency of the SSRS-S Social Skills scale was also adequate ($\alpha = .80$ and $.84$ for females and males, respectively). Test-retest reliability was good for the SSRS-T ($r = .84$ to $.93$), but lower for the SSRS-S ($r = .68$). Studies have also supported the criterion and construct validity of both SSRS forms (Gresham & Elliot, 1990). In this study's sample, internal

consistency was good for the SSRS-S ($\alpha = .90$ pretest and posttest), but lower for the SSRS-T ($\alpha = .62$ pretest and $.69$ posttest).

STRENGTHS AND DIFFICULTIES QUESTIONNAIRE: TEACHER-REPORT VERSION (SDQ-T)

The SDQ (Goodman, 1997) is a brief behavioral screening questionnaire of children's behavior, emotions, and relationships. Teachers completed the teacher version of the SDQ for ages 4–10 (SDQ-T), which consists of 25 items rated on a 3-point response scale. The 25 items comprise five, 5-item subscales: Conduct Problems, Hyperactivity, Emotional Symptoms, Peer Problems, and Prosocial Behavior. Goodman (2001) reports adequate reliability, as assessed by internal consistency (mean $\alpha = .73$), cross-informant correlation ($M = .34$), and retest stability after 4 to 6 months ($M = .62$). In this study's sample, internal consistency was good ($\alpha = .82$ pretest and $.80$ posttest).

AGGRESSION QUESTIONNAIRE (AQ)

The AQ (Buss & Warren, 2000) is a 34-item, student self-report measure of anger and aggression. In addition to a total scale score, the AQ is comprised of five subscales: Physical Aggression, Verbal Aggression, Anger, Hostility, and Indirect Aggression. Each item is rated on a 5-point Likert scale, from *not at all like me* to *completely like me*. Within a normative sample of youth ages 11 to 12, the internal consistency of the total scale was established at $.92$, and test-retest reliability was adequate ($r = .80$). Research also supports the scale's criterion and construct validity (Buss & Warren, 2000). In this study's sample, internal consistency was high ($\alpha = .92$ pretest and posttest).

PREVENTION PROGRAM: URBAN IMPROV (UI)

UI is an interactive theater and educational program designed to serve racially and ethnically diverse inner-city youth in the Boston Public Schools and throughout New England (Freelance Players, 2004; Magis, 2004). UI is an action-oriented, violence prevention program that addresses a variety of themes related to violence and conflict resolution. For a complete description of the basic UI program, please refer to Kisiel et al. (2006). The basic UI program consists of nine weekly sessions that are 75 minutes in length, taking place during the school day in a local theater space (to which students are bussed and accompanied by their teacher). The UI intervention staff includes a director and four actors, all of whom have extensive training in improvisational theater, expressive arts, and youth education. In this evaluation, each troupe was directed by a program originator, but the four actors were new troupe members trained by the program developers. The

fourth-grade curriculum addresses the following themes: friendship, self-esteem, imagination, fear, peer pressure, fairness, violence/conflict resolution, adventure, and family.

The UI TC extends for nine weeks, consists of three parts (improvisational games and read-aloud activities addressing the nine weekly themes, classroom/behavior management strategies, and monitoring and evaluation), and can be incorporated into standard school lesson plans. Part I presents classroom educational activities that address the weekly content themes from standard UI, themes that are highly relevant to youth violence prevention, including bullying, peer pressure, conflict resolution, and self-esteem. Part I utilizes read-aloud activities featuring selections from books and poems, as well as action-oriented games, both of which reinforce the theme of the week. The TC also includes suggestions for class discussion about the texts and games.

Part II incorporates into the classroom behavior management/conflict resolution and positive reinforcement strategies. The conflict resolution strategy is called "Freeze." In response to conflict between students or between a student and the teacher, the teacher will tell the students involved to freeze, in the same manner that UI freezes action at a point of conflict during their improvisational scenes. Then the teacher will ask for suggestions from the class about how to solve the problem. Once several viable suggestions have been offered, the students involved in the conflict will be asked to pick one of the solutions and implement it. Freeze solidifies the problem-solving process modeled by UI by providing students with opportunities to practice this skill to address real-life problems that arise in the classroom. The positive reinforcement strategy, which recognizes and rewards students for demonstrating prosocial behaviors, has teachers develop a Super Actions Wall (SAW) in their classroom. Each week, students are given SAW Awards by the teacher or nominated by another student for demonstrating positive behaviors based on the weekly UI themes. SAW Awards are placed on the Super Actions Wall, and students who receive them earn a small reward or a note home to their parent/guardian from the teacher relaying the child's positive behavior. For example, to reinforce the theme of friendship, students are rewarded for sharing/giving, including classmates in activities/social time, and being polite and helpful with others.

The curriculum was designed to serve either as a supplement to the basic UI program or as a stand-alone violence prevention program for classrooms and schools. The sole difference between the supplemental and stand-alone curricula is that the supplemental curriculum makes reference to the concurrently administered basic UI program in order to link the two program components and remind students of the weekly theme and associated activities being addressed in both components. Students' primary classroom teachers implemented the TC during regular class times. The

total time dedicated to the TC varied slightly by week, averaging approximately two hours. It is important to note that although two hours is lengthy, the curriculum also addresses reading and writing standards required by school systems. For example, the read-aloud activities were designed to address not only the violence-related theme for the week, but also to provide practice in reading skills, and the discussion questions could easily lend themselves to student essays and writing skill practice.

Fidelity to the basic and comprehensive UI protocols was carefully assessed. For UI fidelity, each UI troupe director maintained a weekly checklist of the troupe's adherence to the four core components (song, improvisational skit, structured discussion, and group scenes). To assess teachers' fidelity to the TC, teachers maintained weekly implementation logs.

Procedure

The evaluation of UI consisted of a quasi-experimental, matched control, multi-outcome evaluation with elementary school students in the fourth grade. Classrooms identified a priori by the school district to receive UI and/or the UI TC were matched for school setting, grade, gender, race, learning level, first language, and socioeconomic characteristics with comparison classrooms. All study classrooms were fourth grade, mainstreamed, with English as first language. The protocol consisted of student- and teacher-report questionnaires. Assessment domains and measure selection were informed by our previous study and the literature on youth violence. Evaluation measures were administered at baseline (pretest) and immediately following (posttest) program completion.

Statistical Design

The study was designed to assess the effectiveness of comprehensive UI as a violence prevention initiative for at-risk youth. Specifically, youth exposed to comprehensive UI were compared to controls and to youth exposed to a single component of the comprehensive UI program (basic UI or TC) on aggressive and prosocial behaviors, as well as on scholastic attention and engagement.

The study utilized a quasi-experimental design. Youth participating in the UI program were drawn from pre-existing, intact classrooms previously selected to participate in the program. Control classrooms were selected from within the same schools whenever feasible, or else from different schools in the same district with similar demographic profiles to intervention schools.

Multilevel analysis (HLM) was used to account for the clustering of students within classrooms. This method is appropriate when there are various levels of data nested within each other. In this study, students were

nested within classrooms. Multilevel models explicitly account for the variance and effects at the student and classroom level. To account for changes from pre- to postintervention, the differences between the posttreatment and pretreatment scores (Delta) were used as dependent variables. The independent variables included UI group, TC group, and the interaction between the two interventions. The present study is mainly exploratory in nature to reach tentative, nonconfirmatory conclusions, and therefore, multiple univariate analyses were used in this situation (Huberty & Morris, 1989). To control for baseline differences in study variables, multilevel analysis of covariance (ANCOVA) was used, with baseline score serving as covariate for all outcome measures. Given that groups were found to differ on age and race (see Results), these demographic variables were entered as covariates in analyses. Adding age and race as covariates changed the significance level for only one outcome variable: teacher-reported hyperactivity on the SDQ, which became nonsignificant. All other significant effects remained and are reported without age and race covaried. In addition, raw scores were used on all measures, as the standard scores for these measures differ by gender. Effect sizes were calculated using Cohen's *d* (Cortina & Nouri, 2000). Data were analyzed using SPSS Version 16.0.

To assess whether the TC demonstrates utility as a stand-alone YVP program when the comprehensive program is not administered, univariate planned comparisons were conducted on all measures, focusing on differences between the TC only and the control groups.

Data collection occurred in intact classroom settings during scheduled school times. Strong efforts were made to administer study measures to students absent on the date(s) of data collection, thus limiting the amount of missing data. Nevertheless, a small number of participants had missing data, ranging from single, unanswered items within a measure to omission of multiple measures due to school absences. Within a discrete measure, missing data was accounted for using the rules established for that measure. For example, on the AQ, as the scoring manual indicates (Buss & Warren, 2000), the median value was substituted for the missing item, as long as no more than five item responses were missing overall or two from any scale. To maximize power, participants with excessive incomplete data on any given measure were dropped from analyses requiring that measure, but were included in all analyses for which they had complete data available.

RESULTS

Preliminary Analyses

DEMOGRAPHIC DIFFERENCES

Two-way (UI/No UI group \times TC/No TC group) univariate analysis of variance (ANOVAs) indicated significant group differences in age, $F(1, 252) = 13.97$,

$p < .01$; the group that received the TC was significantly younger than the group that did not. Pearson chi-square analyses indicated significant group differences in race distribution, $\chi^2(6, N = 252) = 34.92, p < .01$, but not gender, $\chi^2(3, N = 260) = 0.43, p > .05$. This difference in race distribution is likely due to the demographic differences among schools and the fact that due to school size, not all schools contributed one classroom per group. See Table 1 for means, standard deviations, and percentages by group.

BASELINE GROUP COMPARISONS

Descriptive statistics are reported in Tables 2 and 3. At baseline, univariate analyses indicated that treatment groups differed on several student- and teacher-report measures. These differences are indicated in Tables 2 and 3. Such differences have been common in our school-based research, and are thought to be related to differences across schools and communities, as well as to the subjective nature of rating forms. Teachers and students may differ in their approaches to rating forms, and the cultures and norms for these behaviors may differ across classrooms, teachers, and communities.

PROGRAM FIDELITY

Adherence to the UI and TC protocols was assessed. For UI fidelity, each UI troupe director maintained a weekly checklist of the troupe's adherence to the four core components (song, improvisational skit, structured discussion, and group scenes). Analysis of UI troupe directors' weekly checklists revealed a fidelity level of 89% (range = 83–100%) across the eight classrooms receiving UI. Analysis of TC teachers' weekly implementation logs indicated an overall fidelity level of 75% (range = 67–81%), indicating adequate fidelity but suggesting that the TC may have been too lengthy to implement in its entirety on a consistent basis across participating classrooms.

Primary Analyses

To test the effectiveness of the comprehensive UI program compared to basic UI, one-way multilevel ANCOVAs were conducted to assess whether the comprehensive UI program had additive effects on student and teacher reports of prosocial behaviors, aggressive or disruptive behaviors, and scholastic attention and engagement. Overall, there were no significant findings for any of the three outcome domains based on student reports. However, significant effects in some domains were found for teacher-report measures.

TABLE 2 Pre- and Posttreatment Means and Standard Deviations for Teacher Measures

Domain/Variable	Group	Pretreatment			Posttreatment		
		<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
Prosocial skills							
SSRS-T general	UI	69	37.30	12.10	67	38.10	13.50
	TC	38	38.40	16.00	43	37.10	15.40
	UI+TC	71	37.40	13.20	70	43.50	11.00
	Control	74	39.90	15.80	76	44.00	15.50
SSRS-T assertion	UI	69	11.60	4.59	67	11.70	4.86
	TC	38	12.40	5.12	43	12.00	5.09
	UI+TC	71	11.00	4.17	70	13.60	4.08
	Control	74	12.40	5.71	76	14.70	5.64
SSRS-T cooperation	UI	69	12.80	4.80	67	13.30	5.13
	TC	38	13.30	6.16	43	13.20	5.45
	UI+TC	71	13.10	5.80	70	14.90	4.56
	Control	74	14.50	5.52	76	15.30	5.29
SSRS-T self-control	UI	69	12.90	5.35	67	13.10	5.67
	TC	38	12.70	6.05	43	11.90	5.76
	UI+TC	71	13.20	5.28	70	15.00	4.67
	Control	74	13.10	6.17	76	14.00	6.12
Aggressive or disruptive behaviors							
SSRS-T externalizing	UI	69	3.32	3.86	69	3.45	3.86
	TC	38	3.16	3.86	43	3.88	4.02
	UI+TC	71	3.42	3.91	71	2.72	3.39
	Control	74	3.40	4.11	76	3.37	3.94
SDQ-T conduct problems	UI	69	1.67	2.32	68	1.68	2.23
	TC	37	2.41	2.98	43	2.42	2.55
	UI+TC	71	1.94	2.55	71	1.27	1.99
	Control	75	1.91	2.43	76	2.09	2.79
Scholastic attention and engagement							
SSRS-T hyperactivity	UI	69	4.52	3.72	69	4.12	3.50
	TC	38	4.82	4.27	43	5.28	4.27
	UI+TC	71	4.59	4.46	71	3.93	4.10
	Control	74	3.50	3.78	76	3.00	3.50
SSRS-T internalizing	UI	69	4.78	3.19	69	4.13	3.07
	TC	38	3.58	3.45	43	4.49	3.30
	UI+TC	71	4.30	2.71	71	3.46	2.55
	Control	74	2.89	3.30	76	2.43	2.99
SDQ-T emotional symptom	UI	69	2.07	2.42	68	2.20	2.23
	TC	37	2.95	3.18	43	2.70	2.59
	UI+TC	71	2.92	2.66	71	1.37	1.97
	Control	75	1.49	2.36	76	1.87	2.62
SDQ-T hyperactivity	UI	69	3.86	3.48	68	3.65	3.03
	TC	37	4.44	3.89	43	4.73	3.44
	UI+TC	71	4.46	3.82	71	3.49	3.41
	Control	75	2.81	3.14	76	2.84	3.17

Note. Data displayed are raw scores for all measures. UI = Urban Improv alone; TC = teacher curriculum alone; UI+TC = comprehensive Urban Improv program. SSRS-T = Social Skills Rating System for teachers. SDQ-T = Strengths and Difficulties Questionnaire for teachers.

TABLE 3 Pre- and Posttreatment Means and Standard Deviations for Student Report Measures

Domain/Variable	Group	Pretreatment			Posttreatment		
		<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
Prosocial behaviors							
SSRS-S total	UI	68	54.90	13.10	69	55.70	12.80
	TC	41	52.20	15.00	43	53.20	14.10
	UI+TC	70	53.90	12.90	71	55.50	12.40
	Control	73	56.40	11.70	75	55.20	12.80
SSRS-S assertion	UI	68	13.40	3.46	69	13.60	2.88
	TC	41	12.50	3.71	43	12.80	3.37
	UI+TC	70	13.80	3.08	71	13.50	3.14
	Control	73	13.80	3.23	75	13.60	3.34
SSRS-S cooperation	UI	68	14.60	3.26	69	14.80	2.99
	TC	41	14.30	4.06	43	14.00	4.20
	UI+TC	70	14.30	3.54	71	14.80	3.23
	Control	73	15.20	2.88	75	14.70	3.35
SSRS-S empathy	UI	68	15.30	3.93	69	15.10	3.80
	TC	41	14.40	4.54	43	14.60	4.33
	UI+TC	70	14.40	3.94	71	15.20	3.54
	Control	73	15.80	3.13	75	15.30	3.68
SSRS-S self-control	UI	68	11.70	4.10	69	12.20	4.02
	TC	41	11.00	4.51	43	11.80	3.94
	UI+TC	70	11.40	4.12	71	12.00	3.94
	Control	73	11.50	4.21	75	11.60	4.09
Anger/Aggression							
Aggression Ques. total	UI	65	82.80	24.90	69	87.60	25.70
	TC	42	91.10	25.30	43	90.00	26.40
	UI+TC	71	93.80	28.20	71	87.90	26.30
	Control	73	85.30	25.30	75	92.20	28.10
Aggression Ques. anger	UI	65	16.20	6.11	69	17.60	6.50
	TC	42	18.70	6.07	43	18.42	6.34
	UI+TC	71	19.90	6.81	71	17.60	6.49
	Control	73	17.60	5.96	75	18.30	6.87
Aggression Ques. physical	UI	65	18.90	8.68	69	19.70	8.01
	TC	42	21.00	8.21	43	20.20	8.44
	UI+TC	71	21.60	9.50	71	19.30	9.42
	Control	73	19.50	8.24	75	21.20	8.54
Aggression Ques. verbal	UI	65	11.70	4.26	69	13.20	4.37
	TC	42	13.30	3.85	43	13.10	4.85
	UI+TC	71	12.60	4.11	71	12.50	4.00
	Control	73	12.40	3.95	75	13.20	4.37

Note. Data displayed are raw scores for all measures. UI = Urban Improv alone; TC = teacher curriculum alone; UI+TC = comprehensive Urban Improv program. SSRS-S = Social Skills Rating System for students. Aggression Ques. = Aggression Questionnaire.

PROSOCIAL BEHAVIORS

We hypothesized that the comprehensive UI program would lead to an increase in prosocial behaviors. Study results indicated that the interaction between UI and TC had a significant effect on the change in several

teacher-reported, prosocial behavior measures. Analyses indicated that the interaction of basic UI and the UI TC had a significant effect on the Social Skills scale of the SSRS-Teacher measure, $F(1, 11) = 8.12, p < .05, d = 0.30$, from pre- to postintervention. Specifically, the comprehensive UI program significantly increased cooperation, $F(1, 11) = 5.25, p < .05, d = 0.35$; assertion, $F(1, 11) = 5.51, p < .05, d = 0.14$; and self-control, $F(1, 11) = 9.60, p < .01, d = 0.29$. A significant effect for the comprehensive UI program was also found on the Prosocial subscale of the SDQ-Teacher, $F(1, 10) = 5.20, p < .05, d = 0.56$.

AGGRESSIVE AND DISRUPTIVE BEHAVIORS

We hypothesized that the comprehensive UI program would act as a preventative measure for the development of aggressive or violent behaviors. However, multilevel analyses demonstrated no significant effects of the basic or comprehensive UI programs in this domain.

SCHOLASTIC ATTENTION AND ENGAGEMENT

We hypothesized that comprehensive UI would enhance students' attention and engagement in the classroom by reducing their internalizing symptoms and hyperactivity. Significant multilevel effects were found for one measure of attention and engagement (SSRS-T). Based on teacher reports on the SSRS, the comprehensive UI program significantly decreased both internalizing problems, $F(1, 11) = 5.92, p < .05, d = -0.35$, and hyperactivity, $F(1, 11) = 4.79, p < .05, d = -0.10$.

Secondary Analyses

To assess whether the UI Teacher Curriculum demonstrates utility as a stand-alone YVP program when the comprehensive program is not administered, univariate planned comparisons were conducted on all measures, focusing on differences between the TC only and the control group. Results were mixed.

PROSOCIAL BEHAVIORS

Simple contrasts on delta scores for student-report measures of prosocial behaviors indicated no differences between the two groups from pre- to posttreatment. Teacher-reports revealed significant differences in the direction opposite of that hypothesized. Specifically, on the SSRS-T and SDQ-T, simple contrasts demonstrated greater gains in cooperation, $F(3, 247) = 3.06, p < .05, d = -0.28$; assertion, $F(3, 247) = 10.37, p < .01, d = -0.68$;

self-control $F(3, 247) = 5.41, p < .01, d = -0.45$; and general prosocial skills, $F(3, 249) = 9.12, p < .01, d = -0.19$, for the control group than the TC only group.

AGGRESSIVE AND DISRUPTIVE BEHAVIORS

Planned comparisons for measures of aggressive behaviors revealed significant differences between the control and TC only groups according to both teacher and student reports. One teacher-report measure (SSRS-T) indicated a greater increase in externalizing problems (SSRS-T) over time in the TC only group compared to the control group, $F(3, 250) = 4.10, p < .01, d = 0.29$. A second teacher-report measure assessing conduct problems (SDQ-T) found the opposite pattern, as disruptive behaviors increased to a greater degree in the control group, $F(3, 249) = 4.59, p < .01, d = -0.07$.

On the AQ, student-ratings on all subscales indicated significant group differences from pre- to post-treatment, with the TC only group demonstrating a decrease in aggression over time, whereas youth in the control group demonstrated an increase in aggression. Specifically, this finding held for the total AQ score, $F(3, 249) = 3.85, p = .01, d = -0.30$, as well as for subscales assessing physical aggression, $F(3, 249) = 2.98, p < .05, d = -0.30$; verbal aggression, $F(3, 249) = 2.67, p < .05, d = -0.20$; and anger, $F(3, 249) = 4.71, p < .01, d = -0.13$.

SCHOLASTIC ATTENTION AND ENGAGEMENT

Simple contrasts performed on one measure of hyperactivity (SDQ-T) but not a second measure of hyperactivity (SSRS-T) revealed a significant group difference, with hyperactivity decreasing over time in the control group but increasing in the TC only group, $F(3, 249) = 5.59, p < .01, d = 0.18$. Significant differences were found on both measures of teacher-reported internalizing symptoms, but in opposite directions. On the SDQ-T, teacher reports of emotional symptoms indicated decreases in the TC only group and increases in the control group over time, $F(3, 249) = 5.60, p < .001, d = -0.30$. In contrast, on the SSRS-T, teacher-reported internalizing symptoms increased in the TC only group but decreased in the control group, $F(3, 250) = 5.90, p < .01, d = 0.61$.

DISCUSSION

Results of this study replicated and extended the findings of our original study of UI (Kisiel et al., 2006). In replicating most of the previous study's findings with blended theater troupes that consisted of new UI staff being directed by a program originator, this study demonstrates that UI is

transferable to new staff. Thus, the positive impacts of UI can be replicated in other urban school systems implementing the UI curriculum.

A comprehensive UI program was also developed, which includes a teacher-led, classroom-based curriculum, in order to incorporate UI into naturalistic classroom environments on an ongoing basis. It was thought that such a curriculum would enable better integration of the program across settings, as well as allow students more opportunities for assimilation of learning and practice of skills introduced in UI. As hypothesized, the comprehensive UI program appeared to demonstrate an additive effect on two of the three domains of behavior assessed, although the results were somewhat inconsistent. Specifically, elementary school students in the comprehensive UI program showed the greatest increases in prosocial skills, including cooperation, assertion, and self-control, as well as the greatest gains in scholastic attention and engagement. However, these gains were not evident on all measures of these domains. Although the finding from the original study that UI prevented new-onset aggression was not replicated, students in the control group did not show the sharp increases in aggression over time that were evident in the original study. Instead, aggression and conduct problems remained generally stable across time in all four groups.

As in the previous study, findings were more robust for teacher measures. Discrepancies between adult- and child-report data are common in child research (e.g., Achenbach, McConaughy, & Howell, 1987; Kisiel & Lyons, 2001). Student self-appraisal at this age is certainly limited, particularly in regard to more subtle indices of change (e.g., engagement, prosocial skills). We would expect that teachers would be more accurate observers of students' behaviors in these realms.

This study also explored whether the UI TC demonstrated utility as a stand-alone YVP program. Such a finding would be important because the TC can be easily disseminated across school systems and integrated into regular classroom activities by classroom teachers, without requiring the specialized staffing and training of the comprehensive UI program. The TC does not require transportation or implementation by a specially trained group of actor/educators, as does the comprehensive UI program. In fact, school personnel from various districts have contacted the primary author of the TC for information about how to access the curriculum for use in their classrooms.

Within the aggression/disruptive behavior domain, with the exception of a single scale, both teachers and students rated students who received the TC as demonstrating significant decreases in aggressive behaviors compared to controls. These findings suggest that both students and teachers felt the TC was helping with conduct and aggression. It is interesting that this was the only significant finding for student-report measures, and it may suggest a more robust influence of the TC on youth attitudes towards

aggression than upon cultivation of prosocial skills or classroom engagement. Specifically, students who participated in a youth violence prevention program within their classroom not only subsequently exhibited less aggressive behavior as reported by their teachers, but also endorsed physical and verbal aggression and anger to be less socially acceptable than non-participating students.

These findings on students' reports of their own aggression suggest that the TC may hold some promise as a stand-alone YVP program, which may merit further exploration in schools without resources to implement the comprehensive program. Such schools may consider implementing the teacher curriculum and taking steps to embed the curriculum within a larger school environment of violence prevention. In this study, teachers were utilizing the curriculum in isolation within their school, with support only from the external research staff. For schools that do adopt the UI teacher curriculum, we have two recommendations. First, adding a teacher consultation group so that teachers implementing the TC can function as a team and support each other in using it would enrich the program. Second, students and teachers who participated in both UI and the TC likely benefited from the sense of being part of a larger program. Thus, if schools are to successfully utilize the TC without the comprehensive, it should be embedded within a more macro school YVP framework and programming in which the violence prevention principles can be practiced within the larger school environment rather than limited to a single classroom.

Study Limitations

Several limitations of this study warrant discussion. First, there were sampling and design limitations, as students were part of intact classrooms, and classrooms were not randomly assigned to group. Instead, group assignments were made partially based on teacher willingness to participate, and whether the school structure could support implementation of UI and/or the TC. The TC stand-alone group had one less classroom, and thus significantly fewer participants, than the other three groups. All of the available classrooms in the six schools were used, and an additional classroom could not be located prior to the start of this study. Furthermore, due to this study's exploratory nature, multiple significance tests were conducted to find the effects on multiple outcomes. To correct possible Type I error using methods such as Bonferroni correction with a large number of significance tests, a larger sample size would be needed than was feasible given the nested nature of student data within classrooms. Second, the reliance on rating forms as the sole type of outcome measure is limited given their subjective nature. The schools involved did not maintain consistent records of behavioral variables, such as disciplinary referrals, so we were limited in our outcome measurement. In addition, ratings by the

teachers who implemented the curriculum may reflect an expectancy effect, and they may have been more invested in demonstrating improvement in their students. Third, the TC can be time consuming to implement (average of two hours per week), which is challenging for teachers who are under pressure to demonstrate gains in their student's reading, writing, and math scores. However, the TC was designed to address not only the content issues related to youth violence, but also to cover specific reading and writing standards (e.g., understanding literature through read-aloud exercises), and thus supports students' academic learning, as well.

Conclusion

As discussed, study findings were mixed. Future research should assess whether embedding the UI TC within a school-wide YVP effort and instituting teacher learning groups to support each other in its use will enhance its effectiveness. In addition, by replicating some of the findings from the previous study with new troupes, this study lends additional support to UI as an effective YVP program for elementary school-aged children. An important next step will be to evaluate whether this theater model is effective when implemented with older youth in middle and high school or as a tertiary prevention model for youth with known exposure to high levels of trauma and violence.

REFERENCES

- Achenbach, T., McConaughy, S., & Howell, C. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*, 213–222.
- Botvin, G. J., Griffin, K. W., & Nichols, T. D. (2006). Preventing youth violence and delinquency through a universal school-based prevention approach. *Prevention Science*, *7*, 403–408.
- Buss, A. H., & Warren, W. L. (2000). *Aggression Questionnaire Manual*. Los Angeles: Western Psychological Services.
- Chen, G. (2007). School disorder and student achievement: A study of New York City elementary schools. *Journal of School Violence*, *6*, 27–43.
- Cortina, J. M., & Nouri, H. (2000). *Effect size for ANOVA designs*. Iowa City, IA: Sage.
- Cross, J. E., Mohajeri-Nelson, N., & Newman-Gonchar, R. (2007). Project LINK: Improving risk and protective factors through comprehensive services. *Journal of School Violence*, *6*, 23–55.
- Edwards, D., Hunt, M. H., Meyers, J., Grogg, K. R., & Jarrett, O. (2005). Acceptability and student outcomes of a violence prevention curriculum. *Journal of Primary Prevention*, *26*, 401–418.

- Eisenbraun, K. D. (2007). Violence in schools: Prevalence, prediction, and prevention. *Aggression and Violent Behavior, 12*, 459–469.
- Farrell, A. D. (2008). The multisite violence prevention project: Impact of a universal school-based violence prevention program on social-cognitive outcomes. *US Prevention Science, 9*, 231–244.
- Freelance Players. (2004). *Urban Improv educational services*. Chatsworth, CA: Aims Multimedia.
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry, 38*, 581–586.
- Goodman, R. (2001). Psychometric properties of the Strengths and Difficulties Questionnaire (SDQ). *Journal of the American Academy of Child and Adolescent Psychiatry, 40*, 1337–1345.
- Greenberg, M. T., Weissberg, R. P., O'Brien, M. U., Zins, J. E., Fredericks, L., Resnik, H. et al. (2003). Enhancing school-based prevention and youth development through coordinated social, emotional, and academic learning. *American Psychologist, 58*, 466–474.
- Gresham, F. M., & Elliot, S. N. (1990). *Social Skills Rating System Manual*. Circle Pines, MN: American Guidance Service.
- Griffin, J. P., Chen, D., Eubanks, A., Brantley, K. M., & Willis, L. A. (2007). A randomized violence prevention trial with comparison: Responses by gender. *Journal of School Violence, 6*, 65–81.
- Hahn, R., Crosby, A., Moscicki, E., Stone, G., & Dahlberg, L. (2007). Effectiveness of universal school-based programs to prevent violent and aggressive behavior. *American Journal of Preventative Medicine, 33*, S114–S129.
- Holt, M. K., Kantor, G. K., & Finkelhor, D. (2009). Parent/child concordance about bullying involvement and family characteristics related to bullying and peer victimization. *Journal of School Violence, 8*, 42–63.
- Huberty, C., & Morris, J. (1989). Multivariate analysis versus multiple univariate analysis. *Psychological Bulletin, 105*, 302–308.
- Hussey, D. L., & Flannery, D. J. (2007). Implementing and evaluating school-based primary prevention programs and the importance of differential effects on outcomes. *Journal of School Violence, 6*, 117–134.
- Kisiel, C., Blaustein, M., Spinazzola, J., Schmidt, C. S., Zucker, M., & van der Kolk, B. (2006). Evaluation of a theater-based youth violence prevention program for elementary school children. *Journal of School Violence, 5*, 19–36.
- Kisiel, C. L., & Lyons, J. S. (2001). Dissociation as a mediator of psychopathology among sexually abused children and adolescents. *American Journal of Psychiatry, 158*, 1034–1039.
- Magis, W. (2004). Urban Improv: A portrait of an educational drama organization. *Youth Theater Journal, 18*, 30–44.
- Margolin, G., & Gordis, E. B. (2004). Children's exposure to violence in the family and community. *Current Directions in Psychological Science, 13*, 152–155.
- Massey, O. T., Boroughs, M., & Armstrong, K. H. (2007). School violence interventions in the Safe Schools/Healthy Students Initiative: Evaluation of two early intervention programs. *Journal of School Violence, 6*, 57–74.

- Nation, M., Crusto, C., Wandersman, A., Kumpfer, K. L., Seybolt, D., Morrissey-Kane, E., et al. (2003). What works in prevention: Principles of effective prevention programs. *American Psychologist, 58*, 449–456.
- Office of Juvenile Justice and Delinquency Prevention (OJJDP). (1995). *Delinquency: Prevention works*. Washington, DC: U.S. Department of Justice.
- Sprague, J., & Walker, H. (2000). Early identification and intervention for youth with antisocial and violent behavior. *Exceptional Children, 66*, 367–379.
- Stein, B. D., Jaycox, L. H., Kataoka, S. H., Wong, M., Tu, W., Elliott, M. N. et al. (2003). A mental health intervention for schoolchildren exposed to violence: A randomized controlled trial. *Journal of the American Medical Association, 290*, 603–611.
- Twemlow, S. W., Fonagy, P., & Sacco, F. C. (2001). Creating a peaceful school learning environment: A controlled study of an elementary school intervention to reduce violence. *American Journal of Psychiatry, 158*, 808–810.
- Weissberg, R. P., Kumpfer, K. L., & Seligman, M. E. P. (2003). Prevention that works for children and youth: An introduction. *American Psychologist, 58*, 425–432.
- Zucker, M., Villaflor, E. B., Holden, J., Jones, R., & Spinazzola, J. (2006). *Urban Improv for the classroom: 4th grade teacher curriculum*. Retrieved January 15, 2009, from www.traumacenter.org
- Zyromski, B. (2007). African American and Latino youth and post-traumatic stress syndrome: Effects on school violence and interventions for school counselors. *Journal of School Violence, 6*, 121–137.