The Potential of Coaching as a Strategy to Improve the Effectiveness of School-Based Substance Use Prevention Curricula

Christopher L. Ringwalt, DrPH
Melinda M. Pankratz, PhD
William B. Hansen, PhD
Linda Dusenbury, PhD
Julia Jackson-Newsom, PhD
Steven M. Giles, PhD
Paul H. Brodish, MSPH

Research-based substance use prevention curricula typically yield small effects when implemented by school teachers under real-world conditions. Using a randomized controlled trial, the authors examined whether expert coaching improves the effectiveness of the All Stars prevention curriculum. Although a positive effect on students' cigarette use was noted, this finding may be attributed to marked baseline differences on this variable across the intervention and control groups. No effects were found on students' alcohol or marijuana use or on any of several variables thought to mediate curriculum effects. The effects of coaching on teachers may not become evident until future years, when they have moved beyond an initial mechanical delivery of the curriculum.

Keywords: *school-based programs; substance use; prevention*

School-based substance use prevention curricula constitute the primary means by which antidrug messages are delivered to adolescents. As of 1999, prevention curricula were implemented in more than 80% of the nation's middle schools (Ringwalt et al.,

---

Christopher L. Ringwalt and Melinda M. Pankratz, Pacific Institute for Research and Evaluation, Chapel Hill, North Carolina. William B. Hansen, Tanglewood Research Foundation, Greensboro, North Carolina. Linda Dusenbury, Tanglewood Research Foundation, Southern Pines, North Carolina. Julia Jackson-Newsom, Tanglewood Research Foundation, Greensboro, North Carolina. Steven M. Giles, Wake Forest University, Department of Communication, Winston-Salem, North Carolina. Paul H. Brodish, Pacific Institute for Research and Evaluation, Chapel Hill, North Carolina.

Address correspondence to Christopher L. Ringwalt, Pacific Institute for Research and Evaluation, 1516 E. Franklin St., Suite 200, Chapel Hill, NC 27514; e-mail: ringwalt@pire.org.

This study was supported by a grant from the National Institute on Drug Abuse (Grant No. 5 R01 DA016098, "Promoting Fidelity Using Remote and Onsite Support"). The authors would like to thank our coaches, Emily Maso, Donna Pittman, and Cicely Wilson, for their hard work as trailblazers during the opportunities and challenges of coaching together. The authors are also indebted to Dr. Inez Drummond of Chicago Public Schools for her support and to all the teachers who collaborated on this study.

Health Education & Behavior, Vol. XX (X): xx-xx (Month XXXX)
DOI: 10.1177/1090198107303311

2002). In recent years, a number of such curricula have been widely promoted as effective by several federal agencies (e.g., National Institute on Drug Abuse, 2003) and private organizations (Drug Strategies, 1999). However, a steadily accumulating body of evidence reveals that even these curricula have only modest effects when implemented under real-world conditions. Effect sizes obtained from large effectiveness trials tend to be substantially less than those obtained by program developers within the context of efficacy trials, which are conducted under optimal conditions (Tobler et al., 2000).

One of the challenges to implementing research-based prevention programs may be the inadequacy of training and ongoing support that teachers receive. This problem is likely to be particularly salient to the extent that the prevention program is a curriculum requiring innovative teaching strategies that are thought to be critical to its success, but with which teachers may be insufficiently familiar or comfortable (Rogers, 1995). Prevention programs typically provide teachers with an initial training workshop of only 1 or 2 days' duration, which may not give them an opportunity to practice the skills and activities required by the program or to provide feedback on its delivery. Indeed, most trainings of this nature are designed to be introductory and thus do not fully develop teachers' mastery of the concepts and skills that are required to achieve the program's objectives (Showers, Joyce, & Bennett, 1987). Ongoing training would thus seem to be necessary to develop teachers' mastery and skills.

Expert coaching is a prime example of the type of ongoing training that supports teachers in educational settings. Coaching, also known as mentoring (McCormick & Brennan, 2001), has been defined as a collaboration between professional colleagues "to reflect upon current practices; expand, refine, and build new skills; share ideas; conduct action research; teach one another; or problem solve" (Robbins, 1995). Coaching has long played an important role in training new teachers (Ralph, 2000) and has demonstrated its value as a means to improve teacher performance in several academic areas (Gingiss, 1993). For example, research by Joyce and Showers (1981) suggested that when teachers were given information about theory, only 50% successfully transferred theory into practice in the classroom. A combination of explaining the theoretical concepts and demonstrating how they translate into practice, followed by opportunities for the teacher to practice skills in the training environment with supportive feedback, increased the effective transfer of the practice into the classroom to 65%. By contrast, of those teachers who received on-site coaching in addition to being provided with theory, a demonstration, and the opportunity to practice with feedback, fully 95% used appropriate skills in the classroom. In addition to increasing teachers' skill utilization, coaching has also been shown to increase teacher career satisfaction (Edwards & Newton, 1994). And perhaps most important, the students of teachers receiving peer coaching have demonstrated improved performance on relevant test measures when compared to students whose teachers were not coached (Showers, 1984).

The reasons why coaching increases teacher performance have yet to be fully explored. There are, however, several explanations that might be considered. First, unlike other limited methods of training that introduce entry-level content or teaching methods (National Implementation Research Network [NIRN], 2004) or invite teachers to consider hypothetical situations, coaching is typically applied to real-life settings (van Driel, Beijaard, & Verloop, 2001). Such settings present teachers with challenges that often cannot be anticipated, much less addressed, within the context of standardized training activities. Advice and suggestions pertinent to specific situations can provide both motivation for the teacher to learn and opportunities to which the coach can respond (Joyce & Showers, 2002). Indeed, some observers have argued that coaching should be considered as a continuous set of activities (Joyce & Showers, 2002). Second,

hands-on coaching and peer feedback assist teachers in viewing their behavior more objectively than when they are left to their own devices (Jones & Nisbett, 1972). As observers, coaches see the classroom from a different physical and psychological vantage point than do teachers, which generates new information about teaching practices that can be shared and tested (Taylor & Koivumaki, 1976). Coaching can thus prompt teacher self-reflection and analysis (Veenman, Gerrits, & Kenter, 1999). Third, coaching begins with an assumption that the coach is a peer who can be trusted to provide meaningful feedback (Mello, 1984). Supportive peers constitute a natural source of motivation, support, and inspiration (McCormick & Brennan, 2001; NIRN, 2004; Spouse, 2001). Finally, coaches provide support for novice teachers as they engage in new interactive strategies with which they may be uncomfortable, particularly when these strategies generate what the teachers perceive to be a negative response from the class (Bierman et al., 2002).

We have found two randomized controlled trials that, although not evaluating the effectiveness of coaching itself, have yielded some tangentially pertinent results concerning the strategy's potential in the field of substance prevention. In the first, Allison, Silverman, and Dignam (1990) compared teachers who were subjected to an intensive staff development effort that included coaching to those who received either in-service training or no training at all. Students of teachers who received intensive staff development were more likely to report that they intended to abstain from alcohol. However, no differences were found in their actual use of alcohol or tobacco or on a variety of other measures.

More recently, Abbott and colleagues (1998) used an experimental design to test whether a series of teacher workshops, which included discussions of classroom management, cooperative learning methods, and interactive teaching, as well as opportunities to practice and receive coaching, improved implementation of program components and student outcomes. The mean implementation score for program components among those teachers attending workshops was double that of the control group that did not receive this training, and students' standardized achievement test scores were positively affected at least 2 years postintervention. Joyce and Showers (2002) have suggested that the primary purpose of coaching is to help teachers implement innovative programs in such a way that direct effects on students can be assessed. This article reports the results of an experimental study designed to test the effect of an initial year of coaching on students exposed to the All Stars middle school substance use prevention curriculum. We hypothesized that students of teachers to whom we provided systematic coaching would manifest improvements in three student substance use outcomes and five key mediating variables targeted by the program. We also examined, as a secondary outcome, the proximal effects of coaching on teachers' behaviors.

METHOD

The All Stars Curriculum

All Stars is an evidence-based prevention program designed to reduce adolescent substance use, sexual behavior, and violence through changes in specific mediating variables (Hansen, 1996; Harrington, Giles, Hoyle, Feeney, & Yungbluth, 2001). Hansen (1992) reviewed key risk and protective factors and found that among the strongest mediators of alcohol, tobacco, and other drug use were normative beliefs,

personal commitments not to use substances, and perceptions that substance use would interfere with personal values and lifestyles, bonding to school, and parental attentiveness. All Stars has been extensively implemented and evaluated and has repeatedly yielded evidence of effectiveness (Hansen & Dusenbury, 2004; Harrington et al., 2001).

The All Stars curriculum can be delivered either by program specialists or regular classroom teachers. The curriculum consists of 24 sessions, of which 14 are required and are administered to the entire class during classroom time. The remaining sessions are considered supplemental and include additional class lessons, small group meetings with peer opinion leaders, and one-on-one meetings between the All Stars facilitators and students. The program includes interactive and cooperative learning activities such as debates, games, and general discussion. Homework is assigned to increase interaction between students and parents and to allow parents to play an active role in the program. Each session is designed to affect at least one of the curriculum's five mediating variables.

Study Sample and Respondents

We conducted this study in 43 middle schools in the Chicago Public School District. Participants were 43 teachers and their respective seventh-grade students. Of these, 21 classes implemented the All Stars curriculum in the spring of 2004, 20 implemented the curriculum in the fall of 2004, and 2 implemented it in the spring of 2005. All schools agreed, as a criterion for recruitment, to designate one staff person to implement the curriculum to at least one class of seventh graders during the course of the academic year. This staff person was most often a regular classroom teacher, but guidance counselors, social workers, physical education teachers, and teaching assistants were also assigned to deliver the program.

Parents of participating students were provided a consent form that explained the study, informed them of our intent to videotape their children's classrooms, and requested their permission to survey their children. To secure parental consent, we both mailed the consent form to the parents of participating teachers' students and sent the forms home with the students themselves. We received parental permission to survey a total of 961 of 1,054 eligible students. Our consent rates for youth in our coached and noncoached groups were 92.5% and 89.4% of the total eligible population of seventh graders, respectively. Student assent was also obtained.

We randomly assigned schools to either receive coaching as part of their implementation of All Stars or to teach All Stars without coaching. We provided All Stars's standard 2 days of in-person training to teachers in both conditions. All participants, regardless of condition, continued to have access to the Web site for All Stars facilitators and to consultation from the lead All Stars trainer on request.

Table 1 displays the characteristics of participating schools, teachers, and students. About four fifths of the students in participating schools were eligible for the federal Free and Reduced Price Lunch Program. The mean age of students participating in our study was just short of 13. They were evenly divided by gender, and about half were African American and one third were Hispanic/Latino. The great majority of the school staff who administered All Stars was female.

Coaching Intervention

We initially selected as our coaches two individuals with experience both in classroom teaching and in teaching of prevention programs. We provided the standard All Stars training to both coaches, who then taught an entire set of lessons while they received intensive

Table 1. Characteristics of Sample Schools, Teachers, and Students by Treatment Group

	Coached Group (n = 23)		Noncoached Group (n = 20)		Total (N = 43)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Schools[a]						
Size (no. of students)	763.5	397.7	716.3	260.6	741.6	337.9
Ethnic composition						
African American	61.0	43.6	62.1	46.3	61.5	44.3
Hispanic/Latino	25.5	36.7	24.5	35.8	25.1	35.9
White	12.5	24.1	10.3	21.2	11.5	22.5
Other	1.0	2.7	3.1	7.5	2.0	5.5
Teachers	(n = 23)		(n = 20)		(N = 43)	
Age	40.4	10.4	40.5	9.4	40.4	9.9
Gender (male): *n*, %	6	26.1	2	10.0	8	18.6
Race/ethnicity: *n*, %						
African American	12	52.2	9	45.0	21	48.8
Hispanic/Latino	2	8.7	1	5.0	3	7.0
White	8	34.8	10	55.6	18	43.9
Other	1	4.4	0	0.0	1	2.4
Classroom teacher	16	69.6	13	65.0	29	67.4
Participating students	(n = 485)		(n = 412)		(N = 897)	
Age	12.5	0.6	12.5	0.6	12.5	0.6
Gender (male): *n*, %	236	48.9	202	49.0	438	48.9
Race/ethnicity: *n*, %						
African American	261	53.8	209	50.7	470	52.4
Hispanic/Latino	142	29.3	115	27.9	257	28.7
White	53	10.9	39	9.5	92	10.3
Other	18	3.7	30	7.3**	48	5.4
Missing	11	2.3	19	4.6	30	3.3

NOTE: Data unavailable for five schools. Significant difference between intervention and control schools.

a. Data derived from the Common Core of Data for the year 2003-2004;

**$p < .05$.

ongoing supervision and support from the All Stars lead trainer. One of our coaches left after the end of the first year and was replaced by a teacher whom we similarly trained. All coaches practiced teaching the curriculum under the supervision of its lead trainer.

In general, coaches served as experts to help teachers improve their understanding of the content and skills required to teach All Stars effectively. As recommended (Showers & Joyce, 1996; Smart et al., 1979), coaches sought to establish a relationship with the teachers to whom they were assigned during initial training. Coaches met with each teacher in the experimental condition on four occasions. They initially met prior to the teachers' implementation of All Stars to set expectations, arrange schedules, and build rapport and then three times during the course of instruction. The three mid-implementation coaching sessions were scheduled to occur before the implementation of Lessons 4, 8, and 11.

Our model of coaching was informed by reviews of both the practice and the research literature (e.g., NIRN, 2004; Spouse, 2001). Our coaching strategy was designed to progress through three phases of mastery. These phases included (1) the development

of basic program implementation mechanics, which included understanding the objectives of the program, following the program manual, and ensuring the appropriate timing of various activities; (2) the development of skills required to implement interactive teaching techniques to engage students; and (3) the development of strategies designed to integrate and infuse the program into other subjects taught and to identify and address the needs of high-risk youth.

We initially intended for the three phases of coaching to be implemented sequentially during the 3 years of our intervention with each teacher. However, once the coaching began, it became clear that this process would require modification to meet individual teachers' needs. As recommended by a number of observers (e.g., NIRN, 2004; Spouse, 2001), we developed an approach to coaching that was both flexible and opportunistic and that was designed to vary with the particular needs and proficiency of each teacher. Thus, coaches exploited opportunities to improve teacher performance and develop "craft knowledge" (Spouse, 2001) that capitalized on their teachers' skills and attitudes. Coaches made every effort to ensure a continued high level of motivation and confidence (Walker, Koroloff, & Schutte, 2002).

To prepare for coaching sessions, each coach viewed videotapes of her assigned teachers' performance for each lesson taught since the prior coaching session. Coaches reviewed the videotapes to determine whether each step of the curriculum was followed. In addition, they assessed teachers' timing, sequence, mechanics, and organization of the lessons as well as the clarity of their communication.

The goals of coaching during this first year of the project and the methods coaches used to achieve them were as follows. First, coaches sought to build a sustainable *relationship* with teachers to facilitate a free exchange of constructive criticism. Thus, coaches were advised to be constructive, to be descriptive rather than personal in providing criticism, and to highlight positive points for teachers. Second, coaches helped teachers develop a level of *comfort* with the curriculum by discussing strategies for more easily delivering the program. All of the teachers in this study were new to the All Stars curriculum, and many had concerns that extended well beyond an understanding of its content. Third, coaches assisted with finding an appropriate *pace* for teaching the program. Given the interactive nature of All Stars, many teachers become mired in discussion and fail to complete their sessions on time. Coaches were instructed to help teachers see topics on which they were spending too much (or too little) time, devise strategies for managing students who tend to dominate conversation, and be more succinct in introducing and concluding sessions. Fourth, coaches helped teachers *prepare* more effectively for each classroom session by ensuring their understanding of the purpose, content, and teaching strategies required for each lesson. Fifth, coaches sought to promote teachers' *self-efficacy*. It was important that teachers not only felt confident in what they were doing but also believed that by following the coach's advice, they could improve their delivery in general and student outcomes in particular. A component of self-efficacy also involved establishing some basic communication strategies in the classroom for engaging students effectively. Coaches taught simple techniques so that teachers could fairly quickly experience a greater level of attention from their students. These techniques included varying their vocal tone (and avoiding monotone), circulating around the classroom when teaching, and expressing enthusiasm about teaching All Stars.

Data Collection

The primary source of data comprised two student surveys completed by each participating student in both our coached and our noncoached samples, both prior to and

following their exposure to the All Stars curriculum. We prepared individual packets for each student that included an assent form and a copy of the questionnaire that was identified by a unique bar code linked to each student's name. Participating teachers collected signed student assent forms and then administered the questionnaires on our behalf to the students in their classes. We instructed teachers to remain at the front of the classroom during survey administration to reinforce their students' perceptions that responses were and would be kept confidential. The questionnaire took an average of 20 minutes to complete, at the end of which students sealed it in a blank envelope we provided and placed it in a box at the front of the class. Altogether, 538 and 394 students responded to the pretest survey in our coached and noncoached schools, respectively. Of these, 491 and 362 students also completed the posttest survey, yielding response rates of 81.7% and 79.9%, respectively. Using the bar codes built into the surveys, we were able to link all posttests to their respective pretests.

The second source of data comprised a self-administered teacher survey mailed at the conclusion of the All Stars curriculum. All 43 teachers responded to the survey.

Survey Instruments

Student survey items pertinent to the study reported here included questions that tapped their recent use of key substances, scales that measured each of the five curriculum mediators, and student demographics (i.e., gender, age, and race and ethnicity). We assessed the first of these with questions concerning whether students had, in the past 30 days, smoked a cigarette, taken more than a sip of an alcoholic beverage, or used marijuana.

We measured five curriculum mediators: lifestyle incongruence, normative beliefs, commitment, bonding to school, and positive parental attentiveness. Survey items assessing these mediators primarily included Likert-type response options that ranged from 1 (*strongly disagree*) to 4 (*strongly agree*), which we converted to a 10-point scale for ease of interpretation. We then adjusted the valence of the items so that higher scores represented more desirable outcomes (e.g., greater bonding to school) and lower scores were less desirable outcomes (e.g., beliefs that most of their peers use alcohol) and averaged responses to the items that constituted each scale to yield an overall mean scale score. At least one half of the items on each scale had to be completed for a mean scale score to be computed. The number of items constituting the scale, alpha coefficients from previous studies using the scale, alpha coefficients from the current pretest survey, and an example item for each mediator scale were as follows.

Lifestyle Incongruence consisted of 11 items (e.g., "I will have a happier life if I stay away from alcohol"). Alpha coefficients for this measure have ranged from .67 (Fearnow-Kenney, Hansen, & McNeal, 2002) to .78 (Harrington et al., 2001), and the scale as currently constituted generally yields alpha coefficients in the high .70s. The pretest alpha coefficient obtained in this study was .76. Normative Beliefs consisted of 12 items (e.g., "My friends think smoking cigarettes is cool"). Alpha coefficients for this scale have ranged from .79 to .90, with a mean alpha of .88 (Harrington et al., 2001). The Normative Beliefs scale also included items that asked students to estimate what proportion of people their age used alcohol, tobacco, marijuana, and inhalants. The pretest alpha coefficient in this study was .82. Commitment consisted of 11 items (e.g., "I have made a decision not to get high by sniffing fumes"). Alpha coefficients for Commitment have averaged .87 with very little observed variance (Harrington et al., 2001). The pretest alpha in this study was .83. School Bonding consisted of 7 items (e.g., "I like the teachers in this school"). This scale generally yields an alpha coefficient of

.80 (Hansen & Dusenbury, 2004; McNeal, Hansen, Harrington, & Giles, 2004). In this study, the pretest alpha was .78. Positive parental attentiveness comprised 7 items (e.g., "My parents often talk with me about things they think are important"). In previous research, this scale has produced an alpha coefficient of between .79 and .81 (Hansen & Dusenbury, 2004). In this study, the pretest alpha was .77.

Teachers responded to a variety of questions designed to tap their implementation of the All Stars curriculum, which included how many lessons they taught (out of a total of 14). We also asked, for each lesson, how much of the lesson they taught, inviting them to indicate one of three response options that ranged from 0 (*none*) to 3 (*all*); as part of this question, we also provided a summary of each lesson as a mnemonic aid. We then averaged responses to the items to yield an overall score for this index. In addition, we asked how long teachers took to prepare for each lesson (five options that ranged from less than 15 minutes to more than 2 hours. Copies of both the teacher and student instruments are available from the first author upon request). We created the former specifically for this study; the psychometric properties of the latter, which has been used extensively in previous studies, are reported above.

Data Analysis

As displayed in Tables 1 and 2, our coached and noncoached groups differed significantly on three student-level variables we measured at baseline. Students in our noncoached group were more likely to comprise races other than African American, Hispanic or Latino, and White and to use both alcohol and tobacco. As a general strategy, we included students' baseline scores for all curriculum mediator and outcome variables in our tests of the effects of coaching on students as assessed at follow-up. We were also concerned that several teacher- or student-level variables might confound the relationship between our intervention and outcome measures. We thus included as controls the race and gender of the school staff who administered All Stars and whether or not they were a regular classroom teacher.

Because students were nested within schools, we needed to control for the potential intraclass correlation of participants to avoid violating assumptions of independence for ordinary least squares (OLS) regression analyses. We therefore used a mixed regression model to account for any potential problems related to nested data (Diez Roux, 2002; Raudenbush & Bryk, 2002). This model took the general form

$$\begin{aligned} Y_{ij} = {} & \gamma_{00} + \gamma_{01} \times \text{CLSS\_TCH}_j + \gamma_{02} \times \text{MALE\_TCH}_j \\ & + \gamma_{03}\text{AFAM\_TCH}_j + \gamma_{04} \times (\text{CONDIT\_2})_{ij} \\ & + \gamma_{10} \times (\text{PREMED})_{ij} + \gamma_{20} \times (\text{MALE\_STD})_{ij} \\ & + \gamma_{30} \times (\text{AFAM\_STD})_{ij} + \mu_{0j} + r_{ij} \end{aligned} \quad (1)$$

In this model, $\gamma_{00},\ldots\gamma_{04}$ represented the Level 2 (i.e., teacher) coefficients, $\gamma_{10},\ldots\gamma_{30}$ the Level 1 (i.e., student) coefficients (equal to $B_{0j},\ldots B_{3j}$), and μ_{0j} and r_{ij} the residuals or random effects at Levels 2 and 1, respectively. Also, CLSS_TCH stood for classroom teacher, MALE_TCH for male teacher, AFAM_TCH for African American teacher, CONDIT_2 for coached teacher, PREMED for pretest score, MALE_STD for male student, and AFAM_STD for African American student. We assumed that our Level 1 coefficients were fixed across Level 2 units, and we employed the restricted maximum likelihood estimation procedure to estimate both fixed and random effects in the model.

Table 2. Student Substance Use and Mediator Scores at Pretest and Posttest by Treatment Group

	Coached Group		Noncoached Group	
	Pretest	Posttest	Pretest[a]	Posttest[b]
30-day substance use (%)				
Alcohol	17.1	21.9	23.9**	25.6
Cigarettes	4.4	5.8	10.6****	11.2***
Marijuana	4.0	5.8	3.7	5.6
Curriculum mediators[c]				
Lifestyle incongruence	7.6	7.6	7.6	7.7
	(1.5)	(1.6)	(1.5)	(1.4)
Normative beliefs	6.7	6.7	6.7	6.8
	(1.5)	(1.4)	(1.7)	(1.4)
Commitment	7.4	7.4	7.2	7.2
	(1.7)	(1.7)	(1.8)	(1.6)
School bonding	6.6	6.5	6.6	6.5
	(1.8)	(2.0)	(2.0)	(1.9)
Parental attentiveness	7.5	7.4	7.4	7.2
	(1.8)	(1.8)	(1.8)	(1.9)

NOTE: Significant difference between intervention and control schools.
a. Asterisks in this column indicate pretest differences between groups.
b. Asterisks in this column indicate posttest differences between groups.
c. Mean scale scores are listed above standard deviations (shown in parentheses). Scale ranges from 0 (*most likely to use substances*) to 10 (*least likely to use substances*).
$^{**}p < .05$. $^{***}p < .01$. $^{****}p < .001$.

To determine the differential effects on students' 30-day use of alcohol, tobacco, and marijuana on those students in our group of coached schools, we dichotomized item responses as 0 (*no use*) or 1 (*use*) and employed a mixed model for binary outcomes, with the outcome variable assuming a Bernoulli distribution; that is, where $Y_{ij} = \log[\text{Prob}(Y_{ij} = 1|\text{Bernoulli})/(1\text{-}P)]$.

Similarly, we used a mixed model for continuous outcomes to address the same question in regards to All Star's five mediating variables. We used a linear mixed model to evaluate the mediator variables, constructing five separate models using each pretest and posttest mediator score in each model, along with the same control variables specified above. We constructed and ran all our models in HLM Version 6.0 Student Edition.

RESULTS

Results of our multilevel models of the effects of coaching on All Stars's ultimate outcomes of interest, adolescent substance use, are shown in Table 3; effects on its mediators are displayed in Table 4. Note the direction of the variables specified in the column headings of both these tables. Negative beta coefficients on the dichotomous 30-day substance use outcomes indicate a more desirable outcome (i.e., less use), whereas positive beta coefficients on the scales that assessed the All Stars mediators indicate a more desirable outcome (i.e., beliefs and attitudes conducive to nonuse). Table 3 reveals that students receiving the curriculum from a coached teacher had a lower rate of escalation of cigarette use than students receiving the curriculum from a noncoached teacher but that there were

Table 3. Results of Mixed Models Relating Student-Level Substance Use to Teacher-Level Predictors

	Alcohol Use		Cigarette Use		Marijuana Use	
Teacher-Level Predictors	Beta	*SE*	Beta	*SE*	Beta	*SE*
Classroom teacher	0.41	0.28	0.63	0.43	0.73	0.37*
Male	0.46	0.33	0.14	0.55	–0.19	0.58
African American	0.30	0.31	–0.03	0.50	0.61	0.54
Coached	–0.17	0.30	–0.80	0.38**	0.02	0.42

*$p < .10$. **$p < .05$.

Table 4. Results of Mixed Models Relating Student-Level Mediators of Substance Use to Teacher-Level Predictors

	Lifestyle Incongruence		Normative Beliefs		Commitment		School Bonding		Parental Attentiveness	
Teacher-Level Predictors	Beta	*SE*	Beta	*SE*	Beta	*SE*	Beta	*SE*	Beta	*SE*
Classroom teacher	–0.10	0.18	–0.47	0.13****	–0.09	0.14	–0.17	0.18	–0.18	0.16
Male	–0.21	0.21	–0.31	0.18*	0.03	0.20	0.03	0.23	–0.11	0.21
African American	–0.38	0.17**	–0.21	0.17	–0.12	0.18	–0.23	0.24	–0.09	0.16
Coached	–0.06	0.17	–0.02	0.12	–0.07	0.15	0.06	0.20	0.10	0.17

*$p < .10$. **$p < .05$. ****$p < .001$.

no intergroup differences in alcohol or marijuana use outcomes. Table 4 suggests that our intervention had no significant effect on any of the curriculum's specified mediators.

Several teacher-level characteristics predicted student posttest scores for marijuana use and a number of mediators. Specifically, students in classrooms in which their teachers administered the curriculum, relative to other types of implementers (e.g., their guidance counselors), were more likely both to increase their marijuana use and to report changes in their normative beliefs congruent with increased substance use, controlling for all other teacher- and student-level predictors in the model. Furthermore, the direction of this finding in regards to teacher type was robust across all the outcomes we examined. That is, teachers who delivered All Stars as *visitors* appeared to be consistently more successful than those who had continual access to their students. Two other teacher-level predictors were noteworthy. Students of African American teachers were significantly less likely than those of teachers of other races to manifest improvements in lifestyle incongruence, and students of male teachers were less likely than those of female teachers to report improvements in normative beliefs.

Our analyses of teachers' self-reported program implementation in our coached as opposed to our noncoached group yielded a few statistically significant differences. Teachers in our coached group (52%) were more likely than those in the noncoached group (14%) to spend at least 30 minutes preparing for each lesson ($p < .05$) and were marginally more likely to teach each component of each lesson (2.6 vs. 2.5 on a scale of 0 to 3, $p < .10$). Both groups of teachers administered about the same number of lessons (13.7 vs. 13.1 out of 14).

DISCUSSION

Improving the effectiveness of programs once they are disseminated is an important goal for the field of prevention. Coaching offers one strategy for potentially improving the likelihood that prevention programs will achieve their intended objectives. This article reports first-year data from a randomized controlled trial that examines the effectiveness of a coached versus a noncoached implementation of the All Stars drug prevention curriculum. Results suggest that relative to teachers who received standard All Stars training only, those who also were coached administered the program in a way that significantly attenuated their students' increase in cigarette use. However, cigarette use at baseline was more than twice as high in the noncoached as in the coached group, and although we controlled for this disparity by including baseline use rates in our analyses, we may have inadequately controlled for this degree of difference.

Even though we found some modest differences between teachers' self-reported behaviors in our coached and noncoached conditions relative to the amount of time that they spent preparing for each lesson and the amount of the lessons they administered, we believe that our findings related to cigarette use should be viewed with caution. We note, first, that we found no intergroup differences in regards to All Stars's mediators or either of the other two substances examined as dependent variables and, second, that our finding related to tobacco could be attributed to the failure of randomization on cigarette use at baseline.

We considered several possible explanations for our disappointing findings in regards to coaching as a strategy to improve the effectiveness of intended student outcomes. We began by examining curriculum effects on substance use among students in our noncoached (or control) group, as displayed in Table 2, and comparing those to effects noted in prior evaluations of All Stars that have also targeted seventh graders. In our study, we found increases over time for alcohol, cigarettes, and marijuana of 1.7%, 0.6%, and 1.9%, respectively. These can then be compared to increases of 3.3%, 1.5%, and 0.3% for students who received All Stars in an earlier randomized controlled trial, relative to 6.8%, 5.4%, and 2.9% for the control group in the same trial (Harrington et al., 2001). These comparisons suggest that the noncoached All Stars intervention in this present study performed as least as well as the earlier All Stars intervention groups on the majority of indicators.

We next considered whether our coaching strategies, which were developed specifically for this project, may not have been sufficiently robust to address the challenges facing our teachers. For example, all our teachers, regardless of intervention assignment, may have been so focused on the mechanics of administering a new and demanding curriculum that their ability to absorb and use advice from our coaches may have been limited. Indeed, this explanation is consistent with the stages of concern model (Gingiss, 1992), which suggests that when teachers begin implementing a program, they are primarily concerned with mechanical issues (i.e., following the basic steps needed to deliver the program) but are not in a position to consider how the curriculum may affect students' behavior. This model also suggests that teachers are not ready to move to higher order levels of delivery—that is, the skillful application of program content—until they have first reached a "routine" level of use with the program. To the extent that teachers' main concern in the initial implementation of All Stars was on mechanical use issues, we should note more of a divergence between those in the coached and noncoached groups in their second and third years of exposure to coaching, when they will be in a position to learn and implement these higher order skills. We also note that in

the study's initial year, we focused our coaching on lessons that teachers had already completed in the hope that their delivery of the remainder of the lessons would improve. We would therefore expect that teachers' improvement in regards to any given lesson would be more likely to occur in their second or third iteration of delivery.

In the second phase of coaching, our coaches will turn their attention to helping teachers improve the quality of their interaction with their students. In this phase, coaches will assist their teachers in developing a style of instruction that will engage students and will emphasize the use of curriculum-driven prompts to promote student involvement. Included in this phase will be coaching to skills for integrating student comments into discussion, managing the classroom, asking probing questions that require students to think more deeply about program content, and generally practicing both verbally and nonverbally immediate communication (e.g., using nonverbal cues that demonstrate active listening vs. focusing on a curriculum manual). In addition, coaches will help teachers understand the underlying concepts of the program by conceptually linking specific activities to intended changes in its five targeted mediators. During the study's third phase, coaches will focus on helping teachers identify and address the needs of high-risk students. For example, we will invite our teachers to focus their attention on students who have poor normative beliefs, who lack personal commitment, or who do not perceive substance use to be incompatible with their ideal futures. During this phase, coaches will also help teachers develop a strategy for infusing program goals and methods throughout the school. We believe that all three phases represent a natural progression from relatively simple to increasingly complex tasks and reflect the natural evolution in teacher concerns with implementing a new innovation.

We were surprised to find that students of African American teachers did not perform as well as those of teachers of other racial and ethnic backgrounds in regards to lifestyle incongruence. In an ex post facto analysis, we added a term to our model concerning whether students were taught by a teacher of the same or a different race, and the significance of this finding only attenuated slightly (to $p < .06$). However, we believe that the potential for this finding is largely uninterpretable because the relationship reported may have been confounded by contextual factors that we did not measure and over which we had no control.

We were also surprised to find that students whose classroom teachers administered All Stars generally manifested poorer outcomes than students taught by other types of school staff (e.g., counselors and physical education teachers). Students taught by their regular teachers were significantly more likely to use marijuana and less likely to report positive changes in their normative beliefs concerning the extent of their peers' use of substances. Furthermore, the direction of classroom teachers' effects on all outcomes measured was uniformly negative, relative to their nonteacher colleagues. This finding runs counter to what Harrington and colleagues (2001) reported in the Kentucky field trial of All Stars, in which classroom teachers produced *more* positive changes than other curriculum implementers.

In considering the likely cause for this finding, we first dismissed the potential for mode effects of the survey administrators. We went to great lengths to assure our students that their responses would be entirely confidential, which included inviting them to seal their completed survey—on which their names did not appear—in an envelope we provided. In addition, even if students surveyed by their classroom teachers had responded in a systematically different fashion than those surveyed by their nonclassroom teachers, we presumably would have been able to account for this potential source of bias when we controlled for students' baseline scores, as both their pretests and

posttests were administered by the same individual. We have also speculated that for reasons we do not understand, our nonclassroom teachers may have possessed greater skills, or rapport with students, than our classroom teachers. In a recent evaluation of Project ALERT taught by external program leaders, St. Pierre, Osgood, Mincemoyer, Kaltreider, and Kauh (2005) commented that some of the teachers and administrators they interviewed believed that students might respond more positively to a curriculum taught by someone whom they considered an outside expert. Some of their respondents also suggested that those students who had a negative relationship with their teachers, and who were also at higher risk of substance use, may have been more responsive to an outside teacher who was unaware of their history and reputation. To date, this question has not received sufficient research attention. Future studies should address whether, and under what circumstances, curriculum administrators who are external to the classroom or even the school may be superior to teachers.

The most substantial limitation to our study concerned the failure of our randomization procedure to yield equivalence across groups in regards to students' cigarette use at baseline. This finding serves as a reminder of the importance of testing, in randomized controlled trials, for equivalence at baseline of key outcome-related variables; the procedure, on occasion, fails to accomplish its objectives. As mentioned earlier, we attempted to control for this problem by including baseline cigarette use in our models. The continued disparities that we found in cigarette use at follow-up, when controlling for baseline use, suggest that our coaching intervention may have been effective. However, our failure to find analogous positive effects for either any of the mediators or the other two substances assessed (i.e., alcohol and tobacco) led us to interpret the tobacco finding cautiously and to attribute it instead to a randomization failure of sufficient magnitude that we were unable to correct for it. We believe that future analyses of the long-term effects of our coaching intervention will help to shed light on whether this attribution is correct. Once teachers have completed their second and third iterations of All Stars, we will reexamine study data to see whether there are consistent changes over time in regards to students' tobacco use and other indicators. By so doing, we avoid the temptation to draw premature conclusions about the efficacy of coaching and, by so doing, commit a Type I error.

Implications for Practice

The utility of coaching to improve the effects of school-based prevention curricula on substance use has not been demonstrated by the results of data from our initial study year. On the other hand, the first phase of coaching was primarily designed to help address teachers' concerns about mechanical issues related to delivering All Stars and to assist them in becoming sufficiently comfortable with the curriculum that they could achieve a routine level of implementation. We conclude that we should not necessarily have expected to see improvements in student-level outcomes until participating teachers gained sufficient familiarity with the curriculum to incorporate coaches' suggestions when they teach it again in the study's second and third years.

Study findings may also shed light on the failure of effectiveness trials to replicate fully the findings of prevention programs evaluated under efficacy conditions. That is, to the extent that teachers of prevention curricula who participate in efficacy trials are familiar with the curricula under scrutiny and were selected for (or had the opportunity to develop) the understanding and skills required, it seems reasonable that they would achieve better student outcomes than teachers engaged in their initial implementation

of the curriculum. If this is the case, full-scale effectiveness trials should give teachers the opportunity to administer a curriculum under scrutiny to at least one group of students before testing its effects on its objectives. Otherwise, prevention science may have set the bar of evidence so high that practitioners are denied support for substance use prevention curricula that have proven their worth in efficacy trials but have failed to do so in effectiveness trials because teachers were insufficiently familiar with them to implement them successfully.

References

Abbott, R. D., O'Donnell, J., Hawkins, J. D., Hill, K. G., Kosterman, R., & Catalano, R. F. (1998). Changing teaching practices to promote achievement and bonding to school. *American Journal of Orthopsychiatry*, *68*(4), 542-552.

Allison, K. R., Silverman, G., & Dignam, C. (1990). Effects on students of teacher training in use of a drug education curriculum. *Journal of Drug Education*, *20*(1), 31-46.

Bierman, K., Coie, J., Dodge, K., Greenberg, M., Lochman, J., McMahon, R., et al. (2002). The implementation of the fast track program: An example of a large-scale prevention science efficacy trial. *Journal of Abnormal Child Psychology*, *30*(1), 1-17.

Diez Roux, A. (2002). A glossary for multilevel analysis. *Journal of Epidemiology and Community Health*, *56*(8), 588-594.

Drug Strategies. (1999). *Making the grade: A guide to school drug prevention programs*. Washington, DC: Author.

Edwards, J. L., & Newton, R. R. (1994). *The effect of cognitive coaching on teacher efficacy and empowerment* (Research Report No. 19941). Unpublished manuscript, Colorado Christian University, Lakewood.

Fearnow-Kenney, M., Hansen, W. B., & McNeal, R. (2002). Comparison of psychosocial influences on substance use in adolescents: Implications for prevention programming. *Journal of Child and Adolescent Substance Use*, *11*(4), 1-24.

Gingiss, P. L. (1992). Enhancing program implementation and maintenance through a multiphase approach to peer-based staff development. *Journal of School Health*, *62*, 161-166.

Gingiss, P. L. (1993). Peer coaching: Building collegial support for using innovative health programs. *Journal of School Health*, *63*(2), 79-85.

Hansen, W. (1992). School based substance abuse prevention: A review of the state-of-the-art in curriculum, 1980-1990. *Health Education Research*, *7*, 403-430.

Hansen, W. (1996). *All Stars Sixth and Seventh Grade Curriculum, Version 1.32, United States Edition: Specialist Format*. Greensboro, NC: Tanglewood Research.

Hansen, W., & Dusenbury, L. (2004). All Stars Plus: A competence and motivation enhancement approach to prevention. *Health Education*, *104*(6), 371-381.

Harrington, N. G., Giles, S. M., Hoyle, R. H., Feeney, G. J., & Yungbluth, S. C. (2001). Evaluation of the All Stars Character Education and Problem Behavior Prevention Program: Effects on mediator and outcome variables for middle school students. *Health Education & Behavior*, *28*(5), 533-546.

Jones, E. E., & Nisbett, R. E. (1972). The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. E. Kannouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 79-94). Morristown, NJ: General Learning Press.

Joyce, B., & Showers, B. (1981). Transfer of training: The contribution of "coaching." *Journal of Education*, *16*(2), 163-172.

Joyce, B., & Showers, B. (2002). *Student achievement through staff development* (3rd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.

McCormick, K., & Brennan, S. (2001). Mentoring the new professional in interdisciplinary early childhood education: The Kentucky teacher internship program. *Topics in Early Childhood Special Education*, *21*(3), 131-149.

McNeal, R. B. J., Hansen, W. B., Harrington, N. G., & Giles, S. M. (2004). How All Stars works: An examination of program effects on mediating variables. *Health Education & Behavior*, *31*(2), 165-178.

Mello, L. T. (1984). *Peer-centered coaching. Teachers helping teachers to improve classroom performance*. Indian Springs, CO: Associates for Human Development.

National Implementation Research Network. (2004). *Implementation research: A synthesis of the literature*. Retrieved August 11, 2005, from http://nirn.fmhi.usf.edu/resources/publications/Monograph

National Institute on Drug Abuse. (2003). *Preventing drug use among children: A research-based guide for parents, educators, and community leaders*. Bethesda, MD: National Institutes of Health.

Ralph, E. G. (2000). Aligning mentorship style with beginning teachers' development: Contextual supervision. *Alberta Journal of Education Research*, *46*(4), 311-326.

Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage.

Ringwalt, C. L., Ennett, S., Vincus, A., Thorne, J., Rohrbach, L. A., & Simons-Rudolph, A. (2002). The prevalence of effective substance use prevention curricula in U.S. middle schools. *Prevention Science*, *3*(4), 257-265.

Robbins, P. (1995). Peer coaching: Quality through collaborative work. In J. Block, S. F. Everson, & T. R. Guskey (Eds.), *School improvement programs: A handbook for educational leaders* (pp. 54-58). New York: Scholastic.

Rogers, E. M. (1995). *Diffusions of innovations*. New York: Free Press.

Showers, B. (1984). *Peer coaching: A strategy for facilitating transfer of training. A CEPM R&D report*. Eugene, OR: Center for Educational Policy and Management.

Showers, B., & Joyce, B. (1996). The evolution of peer coaching. *Educational Leadership*, *53*(6), 12-17.

Showers, B., Joyce, B., & Bennett, B. (1987). Synthesis of research on staff development: A framework for future study and a state-of-the-art analysis. *Educational Leadership*, *45*(3), 77-87.

Smart, D. A., Blase, K. B., Smart, D. I., Graham, K., Collins, S. R., Daly, P. B., et al. (1979). *The teaching-family consultant's handbook* (2nd ed.). Boys Town, NE: Father Flanagan's Boys' Home.

Spouse, J. (2001). Bridging theory and practice in the supervisory relationship: A sociocultural perspective. *Journal of Advanced Nursing*, *33*(4), 512-522.

St. Pierre, T. L., Osgood, D. W., Mincemoyer, C. C., Kaltreider, D. L., & Kauh, T. J. (2005). Results of an independent evaluation of Project ALERT delivered in schools by cooperative extension. *Prevention Science*, *6*(4), 305-317.

Taylor, S. E., & Koivumaki, J. H. (1976). The perception of self and others: Acquaintanceship, affect, and actor-observer differences. *Journal of Personality and Social Psychology*, *33*(4), 403-408.

Tobler, N. S., Roona, M. R., Ochshorn, P., Marshall, D. G., Streke, A. V., & Stackpole, K. M. (2000). School-based adolescent drug prevention programs: 1998 meta-analysis. *Journal of Primary Prevention*, *20*, 275-336.

van Driel, J. H., Beijaard, D., & Verloop, N. (2001). Professional development and reform in science education: The role of teachers' practical knowledge. *Journal of Research in Science Teaching*, *38*(2), 137-158.

Veenman, S., Gerrits, J., & Kenter, J. (1999). Coaching pre-service teachers. In *Abstracts 8th European Conference on Learning and Instruction* (p. 548). Goteborg, Sweden: European Association for Research on Learning and Instruction.

Walker, J., Koroloff, N., & Schutte, K. (2002). *Implementing high-quality team-based individualized services planning: Necessary and sufficient conditions*. Portland, OR: Research and Training Center on Family Support and Children's Mental Health.