

Predicting Implementation Success of an Evidence-based Program to Promote Healthy Relationships among Students Two to Eight Years after Teacher Training

Claire V. Crooks

CAMH Centre for Prevention Science and Western University, London, Ontario

Debbie Chiodo, Suzanne Zwarych, and Ray Hughes

CAMH Centre for Prevention Science, London, Ontario

David A. Wolfe

CAMH Centre for Prevention Science, London, Ontario, and University of Toronto

ABSTRACT

During the past decade, teachers in over 1,500 schools in Canada have been trained to implement the empirically validated Fourth R program, but the sustainability of the program post-training is unknown. For this study, 197 teachers in 26 districts in 6 provinces were surveyed to determine the extent to which they were using the program 2 years or more after training, what modifications they make, and perceived barriers to implementation and sustainability. Results indicated high satisfaction with the program, strong implementation fidelity, and the importance of training, support, and accountability in improving implementation. Training, implementation, and monitoring implications are highlighted.

Keywords: mental health promotion, prevention, school-based programming, implementation, fidelity, sustainability

Claire V. Crooks, CAMH Centre for Prevention Science, London, ON; Western University, London, ON; Debbie Chiodo, CAMH Centre for Prevention Science, London, ON; Suzanne Zwarych, CAMH Centre for Prevention Science, London, ON; Ray Hughes, CAMH Centre for Prevention Science, London, ON; and David A. Wolfe, CAMH Centre for Prevention Science, London, ON; University of Toronto, Toronto, ON.

Funding for this study was provided by the Public Health Agency of Canada to Claire V. Crooks.

Correspondence concerning this article should be addressed to Claire Crooks, CAMH Centre for Prevention Science, 100 Collip Circle, Suite 100, London, Ontario CANADA N6G 4X8; PH: 519-858-5144; FX: 519-858-5149; Email: ccrooks@uwo.ca

RÉSUMÉ

Au cours des 10 dernières années, des enseignants et enseignantes provenant de plus de 1 500 écoles canadiennes ont reçu une formation pour mettre en œuvre le programme Quatrième R (Fourth R), validé par l'expérience, mais on ignore la viabilité du programme une fois la formation terminée. Aux fins de la présente recherche, on a mené un sondage auprès de 197 enseignants et enseignantes provenant de 26 districts au sein de 6 provinces afin de déterminer dans quelle mesure ils utilisaient le programme 2 ans ou plus après avoir suivi la formation, les modifications qu'ils ont apportées et les obstacles perçus à la mise en œuvre et à la viabilité du programme. Les résultats ont révélé un degré élevé de satisfaction par rapport au programme, la fidélité de la mise en œuvre ainsi que l'importance de la formation, du soutien et de la responsabilisation dans l'amélioration de la mise en œuvre. Les répercussions en matière de formation, de mise en œuvre et de surveillance sont soulignées.

Mots clés : promotion de la santé mentale, prévention, interventions en milieu scolaire, mise en œuvre, fidélité, viabilité

There is widespread interest in school-based health promotion and violence prevention for adolescents. Programming based on promoting healthy relationships is especially appropriate for this age group. First, prevention and promotion (or positive youth development) perspectives are highly complementary during adolescence (Catalano, Hawkins, Berglund, Pollard, & Arthur, 2002). Adolescence as a developmental stage presents a window of opportunity for youth to foster healthy relationship patterns as they begin to develop intimate relationships outside the family. It is simultaneously the time during which many health-compromising behaviours emerge, often grounded in these same relationships. Second, the school setting provides many logistical advantages for universal delivery of services. Third, healthy relationships are related to positive mental health and adjustment outcomes (e.g., impact of natural mentoring relationships; DuBois & Silverthorn, 2005), whereas involvement in peer and dating violence is associated with mental health problems such as depression, conduct problems, suicidal behaviours, and substance use (e.g., Ellis, Crooks, & Wolfe, 2009; Silverman, Raj, Mucci, & Hathaway, 2001; Swahn et al., 2008).

Despite these advantages, effective school-based prevention programming has not achieved either widespread or sustained implementation. This article describes the *Fourth R: Skills for Healthy Relationships*, a relationship-based program for adolescents that has been shown to increase healthy relationships and decrease risk behaviours (Wolfe et al., 2009). After a brief description of the program and previous effectiveness data, we turn to issues in implementation and sustainability, based on data from 197 teachers who were trained in the program between two and eight years ago.

The Fourth R is a school-based universal program that promotes healthy relationships and targets peer and dating violence and related risk behaviours. The core grade 9 version of the Fourth R includes a 21-lesson skill-based curriculum that promotes healthy relationships and targets violence and other risk behaviours among adolescents. It is delivered by teachers, who receive one day of training. The program takes a social and emotional learning approach and is based on the contention that relationship skills can be taught in much the same way as the other “three R’s” (Reading, ‘Riting’, and ‘Rithmetic’). The *Fourth R* is composed of three units that address personal safety and violence, substance use, and healthy sexuality/sexual behaviour.

The three units are connected by the thread of healthy relationships and encourage youth to consider the types of relationships they want to have, and not just those they want to avoid. Adolescents receive ample practice role-playing ways to resolve conflict, both as participants and in the role of bystander. In addition to the core program for grade 9 students, there are numerous extensions for other grade levels, curriculum areas, and special populations (see Crooks, Wolfe, Hughes, Jaffe, & Chiodo, 2008, for descriptions).

Effectiveness of the Fourth R

The original grade 9 version of the Fourth R has been rigorously evaluated and demonstrated to improve positive functioning and reduce risk behaviour. Following intervention, students in the intervention group demonstrated increased relationship skills and peer-resistance skills based on observational data rated both by research coders and a group of teachers (Wolfe, Crooks, Chiodo, Hughes, and Ellis, 2012). In addition, post-test results demonstrated that at the school level, the Fourth R program greatly attenuated the relationship between cumulative experiences of child maltreatment and violent delinquency (Crooks, Scott, Wolfe, Chiodo, & Killip, 2007).

The Fourth R was evaluated with a cluster randomized controlled trial involving youth in 20 schools. More than 1,700 students were followed up with 2.5 years post-intervention and found to make healthier choices compared to peers who received health class as usual. Specifically, youth who received the Fourth R instruction reported lower rates of dating violence and higher rates of condom use, with boys showing a more pronounced effect than girls (Wolfe et al., 2009). In addition, the buffering effect for maltreated youth was still present at follow-up with respect to lowering the probability of violent delinquency (Crooks, Scott, Ellis, & Wolfe, 2011). Beyond the effectiveness of the program, teachers find it easy to implement and feel that it provides many benefits for both their students and themselves (Crooks et al., 2008).

On the basis of this evidence, the Fourth R has been identified as an effective practice on the Public Health Agency of Canada's Best Practice Portal (<http://cbpp-pcpe.phac-aspc.gc.ca/intervention/617/view-eng.html>). It is also considered a promising practice on the Substance Abuse and Mental Health Services Administration (SAMHSA) National Registry of Evidence-based Programs and Practices (<http://www.nrepp.samhsa.gov/>). Currently some version of the Fourth R program is used in more than 3,200 schools in North America.

Issues in Implementation, Adaptation, and Sustainability

Although the evidence supporting the Fourth R is strong, having an effective program is only one piece of the puzzle in developing a large-scale mental health promotion effort. For school-based promotion and prevention programs, evidence of efficacy from program developers does not ensure successful implementation (Thaker et al., 2008). There is ample evidence that the extent to which a program is properly implemented translates into program effectiveness. In a review of more than 500 studies evaluated in five meta-analyses, Durlak and DuPre (2008) surmised that the magnitude of mean effect sizes are at least two to three times higher when programs are carefully implemented and free from serious implementation problems. In many cases, even the act of monitoring implementation produces greatly increased effects; for example, in a review of 59 mentoring programs, effect sizes were three times larger in programs that monitored implementation

compared to those that did not (DuBois, Holloway, Valentine, & Cooper, 2002). Clearly both the quality of implementation and the process of monitoring that quality are important.

What are the ingredients for improving ongoing implementation fidelity and sustainability? The ecological framework described by Durlak and DuPre (2008) identifies the importance of considering innovation characteristics, provider characteristics, and characteristics of the prevention delivery system (in this case the schools). Each of these levels has factors that can either increase or impede implementation quality.

At the innovation / intervention level, important characteristics include compatibility, or “the extent to which the intervention fits with an organization’s mission, priorities, and values” (Durlak & DuPre, 2008, p. 337). A program’s fit with the highly structured nature of a school setting is also key, and details such as scheduling can have a significant impact on program success. An evaluation of 432 schools implementing LifeSkills Training in 105 sites found that having longer classes scheduled to implement the program was positively related to both implementation quality and dosage (Mihalic, Fagan, & Argamaso, 2008). Innovation characteristics that have been identified as barriers to implementation include heavy requirements for training, the necessity for small class sizes, onerous student selection processes, and resource-intensive programming (Thaker et al., 2008). The Fourth R was designed with special attention to compatibility and avoids many of the documented pitfalls by aligning with state or provincial curriculum standards and using teachers as implementers within regular class sizes, times, and settings.

Adaptability is another key issue (Durlak & DuPre, 2008), and this has not been studied with the Fourth R. Although the program is adapted to meet provincial or territorial education standards and match local culture, less is known about the extent to which individual teachers modify the program. There is a need to strike a balance between implementation fidelity and appropriate modifications that attend to issues such as ethnicity, gender, social class, culture, developmental level, and the unique needs of at-risk youth (Kerig, Volz, Moeddell, & Cuellar, 2010).

Numerous studies have evaluated the relationships between implementer characteristics and implementation quality. Specifically, the extent to which teachers see the need for the innovation (and see the innovation to be relevant to local needs), the extent to which they think the innovation will achieve the desired benefits, their own self-efficacy about delivering the program, and their skill level have all been found to predict implementation quality (Durlak & DuPre, 2008).

At the level of the delivery system (i.e., school system), numerous researchers have identified the importance of support and accountability from administrators (see Durlak & DuPre, 2008; Fagan & Mihalic, 2003; Mihalic et al., 2008). Teachers are more likely to maintain implementation momentum if they have support from their peers and administrators, as well as continuous reminders about aligning priorities (Langley, Nadeem, Kataoka, Stein, & Jaycox, 2010). Furthermore, they need to know that someone notices whether or not they use a program, and that it matters. Conversely, staff turnover can undermine a program, and in many cases staff change schools or jobs even before the first implementation of the program (Mihalic et al., 2008). Finally, training and technical assistance are important, largely because of the role they play in building capacity at the implementer and delivery-system levels.

In summary, the Fourth R is an effective school-based healthy relationship program for adolescents that has been implemented in numerous sites across Canada and the United States. However, there are many unanswered questions about the trajectory of the program post-training. The current study was undertaken to explore issues of implementation quality, modifications, and predictors of ongoing implementation.

The specific objectives of this study were to determine:

1. To what extent are teachers still using the Fourth R more than two years after training?
2. How does their experience with the Fourth R predict implementation fidelity?
3. What modifications have they made to the program and why?
4. What are the barriers associated with implementation and sustainability?
5. To what extent do teacher perceptions of accountability and support increase implementation fidelity?

METHOD

Participants

A full ethics protocol was submitted and approved by the head of the Research Ethics Board and by the Centre for Addiction and Mental Health.

Boards were selected for inclusion if training records indicated that three or more teachers had been trained in some version of the Fourth R prior to the fall of 2009. These criteria produced a list of 26 boards. All teachers trained in the Fourth R prior to 2009 in one of these boards were eligible to participate. Teachers trained more recently were excluded because the focus was on sustainability beyond the two years of implementation. Teachers could have been trained in the grade 8, grade 9, or Aboriginal Perspectives grade 9 programs. In some cases, teachers were trained in more than one program. Districts spanned several provinces including British Columbia (9 districts), Alberta (4 districts), Saskatchewan (7 districts), Manitoba (2 districts), Ontario (3 districts) and Nova Scotia (1 district). The size of the boards varied greatly, largely due to differences in the extent to which various provinces have undergone amalgamation.

Eligible teachers were identified on the basis of training records maintained by the fourth author (the Fourth R National Education Coordinator). An information letter and survey link was sent by email to all eligible participants. Participants were asked to send a message declining if they were not willing to complete the survey. Interested participants completed the survey online and received a \$10.00 gift certificate if they chose to provide a mailing address in a separate database. Overall, 467 prospective participants were emailed the information letter and survey link, and of those, 37 emails were returned undeliverable, which led to the assumption that these individuals had retired, left the teaching profession, or changed boards. Of the remaining participants, 17 emailed to decline participation, 9 completed the survey online but were excluded either because they were not teachers or because they had been trained too recently, and 207 neither explicitly declined nor participated. The remaining 197 produced usable surveys, although many of these had some questions left blank. The overall response rate (excluding those who were sent the recruitment email in error) was 47%.

Of the final sample, 60% were female, the average teaching experience was 14.3 years ($SD = 8.2$ years), and years of experience with the Fourth R program was 3.6 ($SD = 2.1$). Distribution across provinces included 18.3% from British Columbia ($n=36$), 11.7% from Alberta ($n=23$), 14.2% from Saskatchewan ($n=28$), 1.5% from Manitoba ($n=3$), 49.7% from Ontario ($n=100$), and 4.6% from Nova Scotia ($n=9$). The disproportionate number of participants from Ontario is largely an artifact of one school board (42.6% of the whole sample) in which all of the Fourth R programs have been developed, piloted, and evaluated. Nearly two-thirds of the sample had been trained two to three years ago ($n=123$; 62%), approximately one-third had been trained four or more years ago ($n=59$; 30%), and the remaining 15 participants did not answer the question, many of them commenting that they could not remember.

Measures

Teachers completed an online survey that was expanded from an earlier study (Crooks et al., 2008). The survey included questions about whether the teachers were still using the Fourth R, how much of it they thought they used the last time they implemented it, the extent to which they felt the training had prepared them to teach the program, what year they were trained, and what version(s) they used. In addition, there were several checklists and two scales. A copy of the survey is available from the first author.

Checklists. Several checklists were provided to identify types of modifications, reasons for modifications, implementation challenges, possible sustainability challenges, and factors that might enhance sustainability. These checklists were based on nine years of program feedback and answers on our previous implementation study (Crooks et al., 2008). Each of the checklists also included an opportunity to add comments.

Perceived benefits scale. There were 10 items assessing perceived benefits of the program in terms of its impact on youth (e.g., raise student awareness about healthy relationships), impact for teachers (e.g., improve teachers' skills with facilitating role plays), and impact on school climate (e.g., improve school culture and climate). The questions were worded such that teachers were asked to think about the extent to which the Fourth R conferred specific benefits in comparison to other health programs and materials they have used. Items were rated on a scale from 1 (not at all) to 5 (very much) and the scale had good internal reliability ($\alpha = .91$).

Perceived support and accountability scale. There were 9 items measuring support and accountability for implementing the Fourth R on behalf of the school administration or board (e.g., Are there additional supports in your school for you to implement the Fourth R?; How important is it to your principal that you teach the Fourth R?). These items were rated on a 5-point scale and used to create a total support and accountability scale ($\alpha = .87$).

Procedures

All eligible educators were emailed by the National Education Coordinator and requested to participate in the study. This email invitation included the purpose of the study, the time commitment, the areas to be addressed, and notification of a \$10.00 gift card for participating. Interested participants used the email link to go to an online survey and complete the questions.

RESULTS

Extent to Which Teachers are Still Using the Fourth R Two Years or More After Training

Overall, 72% ($n=142$) of respondents indicated that they implemented the Fourth R in the most recent school year. Of the 55 teachers who either indicated they had not taught it the previous year or did not answer the question, 50 of them provided reasons for not teaching it. The primary reason was that they did not have the opportunity to teach the program: 82% ($n=41$) indicated that they were no longer in the same assignment (i.e., they had changed departments or had been seconded to the board in a different position), and 8% ($n=4$) were on medical or maternity leave and thus did not have the opportunity to teach. Of the remaining few who chose not to implement, reasons included that the program was too long ($n=4$), and that there were challenges sharing one set of materials ($n=1$).

Participants were asked to estimate the percentage of the program that they used during the most recent year they implemented it. Of the 189 teachers who answered the question, the largest group of respondents indicated that they were implementing 81% or more of the program (39.7% of respondents), 24.9% of respondents reported using 61–80%, 17.5% indicated using 41–60%, 12.7% indicated using 21–40%, and the remaining 5.3% indicated that they used less than 20% of the program. Those participants who reported using 81% or more of the program were coded as “high-fidelity implementers” for further analysis, with 77 teachers rating in the high category and 114 in the low category. There were no sex differences between high and low fidelity implementers $\chi^2(1, N=189) = .31, p = .55$.

High- and low-fidelity implementers were compared on a number of demographic and implementation experience variables, including years of teaching experience, years of Fourth R experience, perceived readiness based on training, and perceived positive impact on students, teachers, and school climate. High-fidelity implementers had more years of experience with the Fourth R, felt better prepared to teach the Fourth R after the training, and reported more perceived benefits of the program than low-fidelity implementers (see Table 1). High- and low-fidelity implementers did not differ on years of overall teaching experience.

Table 1
Differences in Teaching and Fourth R Experience between High- and Low-Fidelity Implementers

Variable	High fidelity ($n = 77$) M (SD)	Low fidelity ($n = 114$) M (SD)	F (df)
Years of teaching	13.86 (8.03)	14.74 (8.48)	.51 (1,189)
Years of teaching Fourth R	4.07 (2.21)	3.19 (1.98)	8.16* (1,187)
Satisfaction with training	4.5 (.62)	3.8 (.88)	34.06** (1,188)
Total perceived benefits	51.06 (5.04)	42.13 (6.37)	30.78** (1,182)

* $p < .01$, ** $p < .001$

Modifications to Program

Approximately 86% of respondents indicated that they had modified the program in some way the last time they taught it. Modifications included shortening the program by dropping sessions (34%) or activities (50%). Other modifications included adding new activities (42%), topics (12%), and supplementary resources such as guest speakers and videos (56%). Comments provided by respondents indicated that sometimes the added topics were areas that should have been covered in earlier years but seemed inadequately understood by students (eg., re-teaching anatomy before teaching the healthy sexuality unit). The main reasons provided for modifying the program were time constraints or needs of specific student groups.

Barriers and Successes Related to Implementation and Sustainability

The biggest implementation challenges were timeframes and difficulties with role plays. More minor logistical challenges in resource sharing and AV equipment were reported by a small minority of teachers. Potential sustainability challenges identified by a majority of respondents included the possibility of outdated materials and new teachers who were not trained. Frequency of endorsement of the different barriers is reported in Table 2.

Table 2
Perceived Barriers to Implementation and Sustainability (*N* = 197)

Barriers to implementation	% Identifying barrier
Timeframes difficult to meet	43%
Role plays difficult to carry out	34%
Students resisted role plays	32%
External influences (snow days, assemblies)	29%
Students did not respond well	16%
Not enough training in role plays	14%
Difficulty sharing resources among classes	6%
Difficult to have appropriate AV equipment	5%
Mismatch with local culture	4%
Instructions for some activities unclear	3%
Pressure or resistance from parents	1%
<hr/>	
Potential barriers to sustainability	
Video materials become dated	76%
New teachers who have not received training	56%
Change in provincial standards	45%
New programs get introduced	41%
Pressure from parents	7%

Teachers identified a number of factors that might improve sustainability, shown in Table 3. Keeping materials current, matching emerging technological advances (such as providing Smartboard files), and ongoing training opportunities both for new teachers and as boosters for experienced teachers were the identified priorities.

Table 3
Factors Perceived to Increase Program Sustainability (N = 197)

Factors that would increase sustainability	% Identifying factor
Updated curriculum materials	79%
E-files like Smart Board or Powerpoint	67%
Training for new teachers to use the program	51%
Related professional development opportunities	45%
Booster training for teachers	43%
Financial resources to support program	40%
Support from administrators	17%
Recognition from administrators	12%
Opportunity to be involved in research	11%

The importance of training emerged in a separate question that asked about the likelihood that teachers would attend another Fourth R training if the opportunity were available, to which approximately 60% indicated that they would be likely or very likely to do so.

Perceived Program Benefits

Responses indicated a high level of perceived benefits in all areas, but highest in benefits for students, followed by benefits for teachers, and general classroom/school climate (Table 4).

The Impact of Support and Accountability on Implementation Fidelity

At the univariate level, high-fidelity implementers reported higher levels of support and accountability ($M=38.01$, $SD=7.90$) compared to low-fidelity implementers ($M=27.99$, $SD=9.01$) ($F_{(1,125)}=42.51$, $p<.001$).

Variables that were significantly different between high and low implementers at the univariate level in the areas of experience (Table 1), and the support and accountability scale were entered into a logistic regression to examine the extent to which it was possible to predict types of implementers. Individual experiences were added in a block, followed by systemic influences (i.e., perceived support and accountability). Results of the logistic regression indicated that two factors, satisfaction with the training and perceived support and accountability, contributed uniquely to the prediction of implementation fidelity ($X^2=54.45_{(4,124)}$, $p<.001$). The overall model correctly classified 77% of teachers to their correct group (high- or low-fidelity implementers).

Table 4
Ratings of Perceived Benefits of the Fourth R Compared to Other Health Resources and Programs on a 5-point Scale ($N=186-190$ depending on item)

Perceived benefits for students	M	(SD)
Teaches accurate information about risk behaviours	4.31	(.67)
Raises student awareness about healthy relationships	4.44	(.63)
Teaches students better help-seeking strategies	4.05	(.77)
Encourages students to be more responsible bystanders	4.10	(.77)
Fosters relationship skills among students	4.11	(.76)
Engages students in the activities	4.12	(.83)
Gives students a language to use about relationships	4.07	(.78)
Perceived benefits for teachers		
Provides ideas for teaching strategies to use in other courses	3.92	(.97)
Improves teachers' skills with facilitating role plays	3.67	(.96)
Gives teachers new ideas for activities in other classes	3.93	(.93)
Perceived benefits in classroom and school climate		
Improves relationships in the classroom	3.77	(.83)
Improves school culture and climate	3.48	(.88)

Table 5
Regression Analysis of the Prediction of High- versus Low-Fidelity Implementation from Experience, Support and Accountability

Predictors	B	S.E.	Exp(B)	95% CI
Individual experiences				
Perceived preparedness after training	.93	.36	2.55**	1.26–5.16
Years of experience with the Fourth R	.12	.12	1.12	.88–1.43
Total perceived benefits of program	.08	.04	1.08*	1.00–1.17
Systemic influences				
Perceived support and accountability	.11	.03	1.11**	1.05–1.17

* $p < .05$, ** $p < .01$

DISCUSSION

The results of the survey conducted with nearly 200 teachers trained to implement the Fourth R showed that overall teachers are highly satisfied with the program compared to other available options and that for the most part they continue to implement it, even several years after training. Beyond the benefits they see for students, teachers perceive a benefit to themselves in terms of increasing skills and ideas for other classes, which is noteworthy given that these are substantially experienced educators.

Teachers identified a number of challenges to implementation and sustainability, many of which are being addressed programmatically. For example, Smartboard files are being piloted and additional training supports to address difficulties with role plays are now available, although the impact of these newer strategies remains to be seen. In addition, when the program is updated to match the much-awaited Ontario Grade 9 curriculum standards, the videos will also be dropped, addressing the concern about audiovisual materials becoming dated. Challenges around timeframes and role plays were identified by a significant proportion of respondents, consistent with post-program teacher feedback forms and our earlier implementation survey (Crooks et al., 2008). Finding ways to increase teacher skill and self-efficacy in implementing the role plays is an ongoing area of exploration for our program team.

Implementation fidelity was best predicted by perceived readiness after training and by perceived support and accountability. The importance of high-quality training was particularly salient, given its odds ratio of 2.5 in predicting high-quality implementation. The combination of predictors highlights the need for us to shift from thinking of implementation as an event (i.e., training and starting to use the program) to an ongoing process. The perceived benefits of the program were also a significant predictor of fidelity, but showed low impact in terms of differentiating between the groups. A closer look at item-level statistics showed that while high-fidelity implementers perceived higher rates of benefits, the low-fidelity implementers also rated the program very favourably, resulting in small levels of variance.

LIMITATIONS

Although this study provides a good snapshot of the real-world implementation and sustainability of an efficacious school-based program, there were several limitations. First and foremost, the reliability on self-reported fidelity is suboptimal and susceptible to bias in how teachers recall and rate their implementation. In studies looking at the relationships between different measures of implementation and program outcomes, there is some evidence that observational data are more closely linked to outcomes than are self-report data (e.g., Lillehoj, Griffin, & Spoth, 2004). Similarly, asking teachers about their satisfaction with training retrospectively is not as reliable as a prospective rating following the training but before implementation. Other measurement limitations include the absence of factors such as student characteristics, school policies, and school climate, which are likely to influence implementation (Durlak & DuPre, 2008; Gregory, Henry, Schoeny, & The Metropolitan Area Child Study Research Group, 2007).

Another limitation is the generalizability of the results. Although a response rate in excess of 40% is good for this type of research design, it still means that the findings are based on the experience of fewer than half of the sample. Furthermore, these results can only be generalized to teachers who received training

in the Fourth R. Numerous teachers and schools order the materials without accessing training, and little is known about the implementation of this group. Presumably the challenges in implementation and sustainability would be greatly accentuated among teachers who do not receive the training.

Limitations notwithstanding, several practice implications arise from the current study.

Practice Implications

Successful programs require ongoing updates and innovation and a mechanism to distribute these to existing program users. Teachers' concerns about a program becoming outdated in terms of materials or content highlighted the need for ongoing innovation to meet emerging trends and technological advances. The Fourth R team routinely monitors the field and uses teacher feedback to make these types of revisions. For example, compared to the initial program there is now a lesson on cyber bullying and the availability of e-files. Unfortunately, the revisions are only included in new copies of materials that are printed. What is lacking is a distribution system to inform previous consumers of available updates. As the program moves to more electronic formats and the use of licensing agreements, it will be easier to distribute updates.

Districts need to be encouraged to think of implementation as a process, not an event. The survey findings highlighted the need for districts and administrators to develop an ongoing implementation plan that at a minimum includes regular training to counter retirement and turnover. Booster sessions are another appropriate component given the number of teachers who indicated they would attend additional training if offered. Booster sessions could be used to provide teachers with additional rationale for the program and the opportunity to access a learning community that could help them problem-solve areas of difficulty. Beyond providing additional training opportunities, administrators have an important role in setting the tone of accountability and providing the required supports for those implementing the program (Fagan & Mihalic, 2003).

More comprehensive training and technical support might increase implementation quality. Given that teachers' perceptions of being prepared to implement the program after training more than doubled the likelihood of high-quality implementation, it is essential that a system be in place to identify those who either report low self-efficacy following training or who are observed to have difficulties early in implementation. Other researchers have found that early monitoring and consultation doubled implementation fidelity for teachers implementing a comprehensive literacy program (Greenwood, Tapia, Abbott, & Walton, 2003). One of the few randomized control trials to include different training and technical assistance conditions found that comprehensive technical assistance (composed of on-site coaching and web-based support in addition to the usual workshop) produced better implementation fidelity than the usual workshop alone, although the longer-term impact of this difference is still under investigation (Rohrbach, Gunning, Sun, & Sussman, 2010).

An ongoing monitoring and evaluation plan at the district level would likely increase implementation fidelity. Teachers are more likely to value a violence prevention program if it is valued at the school and board level. The act of monitoring can improve fidelity even beyond the additional information provided by the monitoring (Dubois et al., 2002). Ideally, districts could incorporate ongoing evaluation at the district level to monitor program drift and cessation. This type of internal monitoring and reporting is currently conducted in a large district in western Canada and could serve as a model for other districts. Without such

monitoring, districts can be left with a misplaced confidence that they are using an empirically validated program, when the reality may be that a subset of their teachers is using a subset of the program.

Further research is required to better understand the nature and rationale of modifications made by teachers.

Given the significant proportion of teachers that are modifying the program, there is a need for further research to better understand how modifications are made. There is a recognized need for flexibility, but there is also a big difference between modifications based on running out of time or a lack of confidence in particular areas, and well-conceived adaptations. Each teacher needs to achieve a balance between meeting the unique needs of their students and addressing the goals of the program, and the actual process by which teachers navigate between these tensions is important to understand (Langberg & Smith, 2006).

In summary, where much of the implementation literature focuses on the process during an initial implementation or evaluation year, this study provides an important snapshot of the real-world implementation of an effective program more than two years post-implementation. The results clearly demonstrate that having an effective program that is perceived to be efficacious by teachers is necessary but not sufficient to achieve a sustainable health promotion strategy that is implemented with fidelity. Future directions must focus on the mobilization of the school system in a way to maximize support and accountability for the teachers implementing the program if the full benefits are to be obtained for youth.

REFERENCES

- Catalano, R.F., Hawkins, J.D., Berglund, M.L., Pollard, J.A., & Arthur, M.W. (2002). Prevention science and positive youth development: Competitive or cooperative frameworks? *Journal of Adolescent Health, 31*, 230-239.
- Crooks, C.V., Scott, K., Ellis, W., & Wolfe, D. (2011). Impact of a universal school-based violence prevention program on violent delinquency: Distinctive benefits for youth with maltreatment histories. *Child Abuse and Neglect, 35*, 393-400.
- Crooks, C.V., Scott, K.L., Wolfe, D.A., Chiodo, D., & Killip, S. (2007). Understanding the link between childhood maltreatment and violent delinquency: What do schools have to add? *Child Maltreatment, 12*, 269-280.
- Crooks, C.V., Wolfe, D.A., Hughes, R., Jaffe, P.G., & Chiodo, D. (2008). Development, evaluation and national implementation of a school-based program to reduce violence and related risk behaviors. *Institute for the Prevention of Crime Review, 2*, 109-135.
- DuBois, D.L., Holloway, B.E., Valentine, J.C., & Cooper, H. (2002). Effectiveness of mentoring programs for youth: A meta-analytic review. *American Journal of Community Psychology, 30*, 157-198.
- DuBois, D.L., & Silverthorn, N. (2005). Natural mentoring relationships and adolescent health: Evidence from a national study. *American Journal of Public Health, 95*, 518-524.
- Durlak, J.A., & DuPre, E.P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*, 327-350.
- Ellis, W.E., Crooks, C.V., & Wolfe, D.A. (2009). Relational aggression in peer and dating relationships: Links to psychological and behavioral adjustment. *Social Development, 18*, 253-269.
- Fagan, A.A., & Mihalic, S.F. (2003). Strategies for enhancing the adoption of school-based prevention programs: Lessons learned from the Blueprints for Violence Prevention replication of the Life Skills Training program. *Journal of Community Psychology, 31*, 235-253.
- Greenwood, C.R., Tapia, Y., Abbott, M., & Walton, C. (2003). A building-based case study of evidence-based literacy practices: Implementation, reading behavior, and growth in reading fluency, K-4. *Journal of Special Education, 37*, 95-110.

- Gregory, A., Henry, D.B., Schoeny, M., and The Metropolitan Area Child Study Research Group. (2007). School climate and implementation of a preventive intervention. *American Journal of Community Psychology*, 40, 250-260.
- Kerig, P.A., Volz, A.R., Moeddel, M.A., & Cuellar, R.A. (2010). Implementing dating violence prevention programs with flexibility, fidelity, and sensitivity to diversity: Lessons learned from *Expect Respect*. *Journal of Aggression, Maltreatment, & Trauma*, 19, 661-680.
- Langberg, J.M., & Smith, B.H. (2006). Developing evidence-based interventions for deployment into school settings: A case example highlighting key issues of efficacy and effectiveness. *Evaluation and Program Planning*, 29, 323-334.
- Langley, A.K., Nadeem, E., Kataoka, S.H., Stein, B.D., & Jaycox, L.H. (2010). Evidence-based mental health programs in schools: Barriers and facilitators of successful implementation. *School Mental Health*, 2, 105-113.
- Lillehoj, C.J.G., Griffin, K.W., & Spoth, R. (2004). Program provider and observer ratings of school-based preventive intervention implementation: Agreement and relation to youth outcomes. *Health Education & Behavior*, 31, 242-257.
- Mihalic, S.F., Fagan, A.A., & Argamaso, S. (2008). Implementing the LifeSkills Training drug prevention program: Factors related to implementation fidelity. *Implementation Science*, 3, 1-16.
- Rohrbach, L.A., Gunning, M., Sun, P., & Sussman, S. (2010). The project Towards No Drug Abuse (TND) dissemination trial: Implementation fidelity and immediate outcomes. *Prevention Science*, 11, 77-88.
- Silverman, J.G., Raj, A., Mucci, L.A., & Hathaway, J.E. (2001). Dating violence against adolescent girls and associated substance use, unhealthy weight control, sexual risk behavior, pregnancy, and suicidality. *JAMA*, 286, 572-579.
- Swahn, M.H., Simon, T.R., Hertz, M.F., Ariad, I., Bossarte, R.M., Ross, J.G., Fross, L.A., Iachan, R., & Hamburger, M.E. (2008). Linking dating violence, peer violence, and suicidal behaviors among high-risk youth. *American Journal of Preventive Medicine*, 34, 30-38.
- Thaker, S., Steckler, A., Sanchez, V., Khatapoush, S., Rose, J., & Hallfors, D.D. (2008). Program characteristics and organizational factors affecting the implementation of a school-based indicated prevention program. *Health Education Research*, 23, 238-248.
- Wolfe, D.A., Crooks, C.V., Chiodo, D., Hughes, R., & Ellis, W. (2012). Observations of adolescent peer resistance skills following a classroom-based healthy relationship program: A post-intervention comparison. *Prevention Science*, 13, 196-205. DOI: 10.1007/s11121-011-0256-z
- Wolfe, D.A., Crooks, C.V., Jaffe, P.G., Chiodo, D., Hughes, R., Ellis, W., Stitt, L., & Donner, A. (2009). A universal school-based program to prevent adolescent dating violence: A cluster randomized trial. *Archives of Pediatric and Adolescent Medicine*, 163, 693-699.