Quality of implementation: developing measures crucial to understanding the diffusion of preventive interventions

Linda Dusenbury¹,⁴, Rosalind Brannigan², William B. Hansen¹, John Walsh³ and Mathea Falco²

Abstract

As prevention programs become disseminated, the most serious threat to effectiveness is maintaining the quality of implementation intended by the developers. This paper proposes a methodology for measuring quality of implementation in school settings and presents data from a pilot study designed to test several of the proposed components. These methods included assessments of adherence, quality of process, the positive or negative valence of adaptations, teachers’ attitudes and teachers’ understanding of program content. This study was conducted with 11 teachers who had varying degrees of experience who taught Life Skills Training. Observation and interview data were collected during visits to schools. Results suggest that quality of implementation can be measured through observation and interview. Teachers varied in adherence and quality of program delivery. All teachers made adaptations to the program. Experienced teachers were more likely to adhere to the curriculum, deliver it in a way that was more interactive and engaging to students, communicate the goals and objectives better, and make positive adaptations. The field can use these findings as the basis for exploring strategies for measuring and improving quality of implementation.

Introduction

The field of drug prevention must learn how to achieve effectiveness when programs are disseminated. Other fields, such as education, treatment and prevention more broadly defined, are also in need of this understanding. Widely disseminated programs like DARE (Rosenbaum et al., 1994; Clayton et al., 1996) and Here’s Looking at You, 2000 (Hopkins et al., 1988) have been evaluated by independent investigators in school settings, and have failed to demonstrate effectiveness. In contrast, effective programs have been evaluated by their developers in carefully controlled settings. How these programs perform when disseminated has not yet been documented.

A key variable thought to influence effectiveness in school settings is quality of implementation, which research has shown varies considerably (Dusenbury et al., 2003). High-quality implementation is expected to ensure that research-based programs will be effective when they are disseminated. From what is currently known or theorized about achieving quality of implementation (Backer, 2001; Dusenbury et al., 2003), the following variables are likely to be important: (1) dosage—providing sufficient exposure to the program, (2) adherence—following program methods and completing its delivery as outlined in a manual or curriculum guide, (3) quality of process—engaging students through their active participation, and (4) adaptation—modifying the program to meet developmental and cultural needs. In addition, other issues that

¹Tanglewood Research, 7017 Albert Pick Road, Suite D, Greensboro, NC 27409, ²Drug Strategies, 1755 Massachusetts Avenue, Room 821, Washington, DC 20036 and ³Washington Office on Latin America, 1630 Connecticut Avenue, NW Suite 200, Washington DC 20009, USA
⁴Correspondence to: L. Dusenbury; E-mail: lindadusenbury@tanglewood.net

doi:10.1093/her/cyg134
may explain a program’s success or failure include: (5) teachers’ attitudes about a program, (6) teachers’ understanding of the concepts being addressed and (7) teachers’ prior experience.

To this point, only dosage and adherence have been systematically studied. Research in controlled settings shows these variables to predict student outcomes (Rohrbach et al., 1993; Botvin et al., 1995). In contrast, little is known about the quality of process in school settings.

A methodology for examining the role of quality of implementation under real-world conditions has not yet been developed. Recently, two studies (Hallfors and Godette, 2002; Ringwalt et al., 2002) introduced a methodology for studying quality of implementation that relied on teachers’ self-reports and did not include classroom observations. Observations may be crucial because teachers are known to be biased in their reports (Hansen and McNeal, 1999).

The purpose of this study was to examine quality of implementation in dissemination using an observational methodology. This study focused on adherence, quality of process, valence of adaptation, teacher attitudes, teacher understanding of concepts and teacher’s prior experience with prevention. It was beyond the scope of this study to measure dosage. We further explored relationships among quality of implementation variables. The study focused on Life Skills Training (LST) (Botvin et al., 1995). LST is the drug abuse prevention program of choice among approximately one in four (26.8%) schools that have adopted a research-based program (Ringwalt et al., 2002). Baltimore, the setting in which this study was conducted, has implemented LST in all middle schools since 1998. Because multi-year, district-wide implementation of research-based programs is rare, this serves as an important example of what happens to research-based prevention in practice.

**Method**

**Subjects and setting**

Thirteen middle schools in Baltimore were approached to participate in this project. Eleven teachers from seven schools participated. Median enrollment in these schools was 769; 89% were African-American, 9% were white, 81% of students were eligible for reduced cost lunches. On average, there were 17 students per observed class.

Participating health education teachers had between 1 and 15 years of classroom experience with a median of 5 years teaching experience overall.

**Procedures**

Teachers had already completed LST and were asked to re-teach one session for observation. Eleven different sessions were observed: three on information, one about media influences, and seven on personal and interpersonal skills. Two trained masters-level research assistants completed observations and conducted interviews. Two sessions were observed by both research assistants. Raw data were subsequently reviewed and scored by two PhD-level researchers.

**Measures**

Observations assessed adherence, quality of process and engagement, adaptation, and teachers’ attitudes. Ratings were supplemented with written notes. Interviews assessed teachers’ attitudes and prior experience with prevention.

**Adherence**

Six items measured adherence. Specifically, observers coded the number of objectives and, separately, major points completed by teachers. Full points were awarded when objectives or major points were met. Half points were awarded when these were partially met. They also provided a summary judgment about the proportion of objectives (on a five-point scale) and major points (on a seven-point scale) that were covered; Cohen’s $\kappa$ was 0.857 on these ratings in jointly viewed classrooms. Observers assessed whether necessary materials were present. They rated and documented evidence of planning by the teacher.

**Quality of process**

Observers rated how well lessons were delivered and received. Ratings were obtained for: (1)
teacher–student interactivity, (2) teacher enthusiasm, (3) teachers’ communication of goals and objectives, (4) student engagement, (5) student attentiveness, and (6) students expressing their opinions.

Valence of adaptation
Observers noted how activities were altered from those outlined in the manual. Researchers subsequently assessed the number of changes and rated whether these were consistent with or detracted from the program’s objectives. The scale ranged from $-2$ to $+2$ with negative scores representing detracting adaptations and positive scores representing enhancing adaptations. The number of and average valence of adaptations were calculated. In interviews, teachers were asked whether and how they had altered LST when they taught it. Researchers counted how many adaptations teachers reported.

Teacher attitudes
Teachers discussed the program’s strengths and weaknesses and assessed the appropriateness of the curriculum for their students. They also reported their satisfaction with the program. Researchers counted the number of positive and negative statements teachers made about the program. Teachers’ attitude scores were created by subtracting the number of negative statements from the number of positive statements.

Teachers’ understanding of concepts
Observers rated how knowledgeable the teacher appeared to be about concepts being taught in the session. Specifically, observers assessed whether teachers seemed to understand the concepts they were presenting, and whether teachers answered students’ questions and guided discussions in desired directions.

Teachers’ prior experience with prevention
Teachers were asked how long they had taught LST and how long they had been teaching other prevention curricula.

---

Results

Measures of quality of implementation

Adherence
Teachers implemented an average of 65% of teaching objectives and 58% of main points. The range for completion was from 45 to 100% for teaching objectives and 30 to 93% for main points covered. All teachers appeared prepared to teach the LST lessons. These six adherence items were standardized and combined into a single scale. Five items contributed to this scale, which had an $\alpha$ coefficient of 0.89.

Quality of process
Observers rated 82% of the sessions as being very interactive. All but three teachers (73%) were rated as delivering the curriculum with great enthusiasm. Observers average rating of teachers’ communication of goals and objectives was 4.4 on a seven-point scale. Nearly half (45%) of sessions were rated as being very engaging. Raters average rating on student attentiveness was 4.6 on a seven-point scale; 45% were given ratings of 6. All but two teachers (82%) were given the highest rating for students’ sharing of opinions. A quality of process scale was created by standardizing items and averaging ratings ($\alpha = 0.92$).

Valence of adaptation
Based on observations, all teachers altered the LST curriculum. The average teacher made 3.5 definable alterations. There were more negative than positive adaptations; 63% of adaptations were judged by researchers to be negative.

There are noteworthy examples of adaptations. One teacher had students role-play a situation instead of discuss it. Conversely, one teacher had students estimate peer marijuana use, but never compared estimates to actual rates, leaving the erroneous impression that marijuana use was commonplace. Observers noted that additions took time away from prescribed activities.

Ten teachers stated they added materials to lessons including examples to make lessons
culturally relevant and added reading materials, videos, testimonials from drug addicts and puppets. Surprisingly, the number of adaptations listed by teachers was negatively correlated \( (r = -0.41; P = 0.22) \) with observers’ counts.

**Teacher attitudes**

All teachers liked teaching LST and thought it to be appropriate for their students. Teachers mentioned an average of 1.9 positive qualities and 1.9 negative qualities. Eight teachers noted that LST provided skills that students do not receive elsewhere. Five commented positively about the manual, training and student workbooks and two noted the program was fun.

On the negative side, more than half (six) said that students found the program boring. Teachers also complained about the need for cultural adaptation (five), insufficient information about alcohol and other drugs (five), time requirements (three said it took too much time, one wanted more activities) and lack of parental involvement (one).

**Understanding of concepts**

Observers’ average ratings of teachers’ understanding was 4.7 on a seven-point scale.

**Teacher’s Prior Experience**

Teachers had taught LST an average of 3.2 years. Teachers had an average 6.3 years experience with prevention, which was highly correlated with number of years teaching LST \( (r = 0.838; P < 0.01) \).

**Relationships among variables**

Correlations were used to explore relationships among the measures (see Table I). It should be noted that with small samples, the magnitude of the relationship between variables can appear inflated.

Teachers’ understanding of LST was strongly correlated with adherence \( (r = 0.784; P < 0.01) \), as was the quality of process scale \( (r = 0.663; P = 0.03) \). Teachers who had taught more years of any prevention program were most adherent \( (r = 0.630; P = 0.05) \), and more likely to meet objectives \( (r = 0.590; P = 0.06) \) and cover major points \( (r = 0.756; P < 0.01) \).

Teachers who made positive adaptations also tended to adhere more to the curriculum, whereas teachers who made negative adaptations generally failed to follow its directives \( (r = 0.474; P = 0.14) \). Classes were more likely to be engaged in high-quality interactions when teachers had more general experience teaching prevention \( (r = 0.593; P = 0.07) \). Teachers who had more years teaching LST \( (r = 0.593; P = 0.07) \) and more years teaching prevention in general \( (r = 0.577; P = 0.08) \) were more likely to make positive adaptations. Overall, more experienced teachers tended to have more negative attitudes towards LST \( (r = -0.475; P = 0.17) \).

**Discussion**

Rogers’ Diffusion of Innovation Theory (Rogers, 1995) focuses on the processes of adoption, implementation, adaptation and institutionalization; however, the field is now more concerned with understanding how to achieve effectiveness under real-world conditions. This study presents several methodological advances that can be used to assess quality of implementation. The findings also suggest strategies for improving implementation in school settings.

The study involved a small number of teachers. Further, teachers were asked to re-teach the sessions observed during this study and might therefore have acted differently than they would under normal conditions. This study therefore should be considered exploratory; nevertheless, the results of this approach had good internal consistency when scales were constructed. Further, statistically significant relationships were observed, suggesting these variables assessed characteristics worthy of further investigation. The methodology relied on observation and interview, not self-report surveys, which may be biased. Ultimately, these variables need to be tested as mediators of program effectiveness.

Even with the small sample of teachers, quality of implementation varied considerably. No teacher carried out lessons exactly as written. The level of
implementation was comparable to what has been observed in rigorously controlled field studies of LST (Botvin et al., 1995). Given the frequency with which adaptations are observed in research and practice, program developers need to anticipate how and when teachers will modify programs and develop guidelines and recommendations to ensure program goals are met.

Teaching experience played an important role in determining the quality of implementation in this study. Further, seasoned teachers were more knowledgeable about the program and better able to communicate its objectives. Among researchers, a commonly held hypothesis is that teachers with more experience are more likely to deviate from prescribed instruction. This hypothesis, not supported by this study, may have been based on experienced teachers’ criticisms of programs more than on the quality of their implementation. We found that even though experienced teachers were more critical of LST they implemented it more completely and with better quality.

The finding that teachers’ self-reports about adaptations were negatively correlated to observations raises serious questions about the validity of these instruments and warrants further study. Because observers are usually thought to be more reliable (Hansen and McNeal, 1999), this raises questions about whether teachers can be expected to accurately report whether they adapt curricula. They may simply not recognize when they are making changes to the curriculum.

Teachers in this study received minimal training; none received formal coaching or technical assistance. Developing methods for accelerating understanding and experience should be a national priority. Teachers at all levels, but especially new teachers, need more training, technical assistance and feedback regarding their performance.

Future research is needed to refine and replicate the methods developed in this study with larger groups of teachers and with other programs. Research is also needed to identify effective strategies for promoting adherence, promoting positive adaptations and achieving high-quality implementation. Ultimately, the field can use these findings as the basis for exploring strategies to improve the quality of implementation.

Acknowledgements

This research was supported in part by grants from the Abell Foundation and the National Institute on Drug Abuse (grant 1R01DA016098).

References


Quality of implementation


Received on December 17, 2003; accepted on September 21, 2004