
Evidence-based practice in school substance use prevention: fidelity of implementation under real-world conditions

S. T. Ennett^{1*}, S. Haws¹, C. L. Ringwalt², A. A. Vincus², S. Hanley^{1,2},
J. M. Bowling¹ and L. A. Rohrbach³

¹Department of Health Behavior and Health Education, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA, ²Pacific Institute for Research and Evaluation, Chapel Hill, NC 27514, USA and ³Department of Preventive Medicine, University of Southern California, Los Angeles, CA 90089, USA

*Correspondence to: S. T. Ennett. E-mail: sennett@email.unc.edu

Received on July 14, 2010; accepted on February 3, 2011

Abstract

Fidelity of program implementation under real-world conditions is a critical issue in the dissemination of evidence-based school substance use prevention curricula. Program effects are diminished when programs are implemented with poor fidelity. We assessed five domains of fidelity—adherence, exposure (dosage), quality of delivery, participant responsiveness and program differentiation (lack of contamination from other programs)—in a subset of respondents ($N = 342$) from a national random sample of public schools with middle school grades ($N = 1721$). Respondents taught 1 of 10 evidence-based universal substance use prevention programs as their primary program during the 2004–05 school year. Their responses to survey questions about their recent implementation practices indicated that fidelity was high for quality of delivery and participant responsiveness, low for program differentiation and modest for adherence and exposure—the two core domains of fidelity. Results suggest the need for continued emphasis on fidelity in program materials, trainings and on-going technical support. Particular attention should be paid to supporting use of interactive delivery strategies.

A sizeable number of school-based substance use prevention programs with demonstrated effects on

youth alcohol, tobacco and other drug use in research trials are packaged for dissemination. School adoption of evidence-based programs has been aided by consumer information available on the Substance Abuse and Mental Health Services Administration's web-based National Registry of Evidence-based Programs and Practices (NREPP) [1] and by federal education policies that promote use of evidence-based prevention programs [2, 3]. School adoption is impressive with almost half of US middle grade public schools using an evidence-based substance use prevention program [4].

With widespread adoption, come questions about how schools are implementing programs. The promise of public health impact on the prevalence of youth substance use when evidence-based programs are transferred to real-world settings depends on the extent to which they are implemented as the program developers intended [5–9]. Program effects are diminished when programs are implemented with poor fidelity [7, 8, 10].

Our purpose is to assess the fidelity of implementation of evidence-based school substance use prevention curricula taught by middle school teachers or other school prevention staff who were using the curricula under real-world conditions, not because they were participating in research. We know relatively little about fidelity of implementation in a non-research context; most fidelity research has been conducted in the context of program evaluations. Our research has implications both for forecasting

the likely effects on youth substance use of school adoption of evidence-based programs and for uncovering aspects of program delivery that may compromise fidelity under natural as opposed to research circumstances.

Definitions of fidelity are variable. Dane and Schneider [11] provided perhaps the most comprehensive schema in defining five domains of fidelity reflected in the prevention program evaluation literature; the schema has been applied to substance use prevention programs [7, 12]. The domains are adherence, exposure, quality of delivery, participant responsiveness and program differentiation.

‘Adherence’ and ‘exposure’ are the core domains in that they measure the extent to which specified program components are delivered as prescribed and the quantity of the program delivered (i.e. dosage). Applied to school-based drug use prevention curricula, adherence encompasses two subdomains: the delivery of specified program ‘content’ and use of specified ‘delivery strategies’ [13, 14]. Both are necessary to achieve effects on youth drug use [13, 15]. Exposure is typically indicated by the number of lessons taught but can reflect combinations of the number of lessons, amount of each lesson covered and adherence to the prescribed schedule.

‘Quality of delivery’ is defined as the aspects of program implementation not directly related to prescribed content and delivery strategies, such as teachers’ enthusiasm, preparedness and attitudes toward the program. The assumption is that teachers who are better prepared and more comfortable with a program’s prescribed methods and who more strongly support its purpose and methods will implement it in a more competent manner [16].

‘Participant responsiveness’ refers to program recipients’ levels of participation and enthusiasm. Participants’ reaction to a given program may be an indicator of the provider’s skill in implementing the program as intended [16]. Process evaluators assert that how the program is delivered, which depends on the program provider, is not the same as how the program is received, which is a function of the target audience [17, 18]. The extent to which participants actively engage with the program bears on its potential effects.

‘Program differentiation’ refers to the absence of contamination from another program that could account for any effects noted. In the research context, differentiation refers to a manipulation check to ensure that participants in the experimental condition received only the planned intervention. In the school drug use prevention literature, program differentiation has been interpreted to mean the extent to which the effects of program components can be differentiated [7, 19]. More consistent with Dane and Schneider’s [11] definition, however, is the possibility of program contamination through simultaneous exposure to other substance use prevention programs. Fidelity may be compromised when a program is altered by the incorporation of materials from another program.

Research evidence suggests substantial variability in the fidelity of implementation of school substance use prevention curricula. The variability may be due in part to the study design, the source of the measures and the domains of fidelity assessed. Fidelity may be higher in efficacy trials where specialists implement the curricula [20] than in effectiveness trials where teachers are typically the providers [16, 21]. One recent study suggests that fidelity may be highest in dissemination research where teachers receive on-going support and technical assistance in addition to training [8, 22]. Fidelity ratings also are typically higher when based on self-reports than on observations by outsiders; observational data are assumed to be more valid than self-reports because the latter are more subject to social desirability bias [23–25].

Findings concerning exposure have been most commonly reported [6]. Several studies suggest that averaged across schools, teachers typically deliver from two-thirds to three-quarters of a curriculum [9, 16, 21, 26], although average estimates as high as 86% have been reported [22]. Fewer studies have measured adherence or quality of delivery. Findings suggest, however, that teachers may achieve higher fidelity on the adherence subdomain of content than on delivery strategies [12, 19, 27, 28]. Several studies have reported favorable estimates of student responsiveness, based on either student or teacher reports [12, 29, 30].

The few studies of real-world implementation, where fidelity was not assessed in the context of research on particular prevention curricula, suggest poor fidelity [14, 31–33]. Hallfors and Godette [32] estimated that teachers in as few as 19% of schools in a relatively large sample of school districts in 12 states were implementing an evidence-based curriculum with fidelity. In a national study of substance use prevention practices in middle schools, teachers were more likely to show better fidelity in adherence to program content than to delivery strategies with only 17% using prescribed interactive delivery strategies [14]. The same study found that the practice of implementing evidence-based curricula in tandem with other programs is widespread [34], suggesting the likelihood of contamination by other programs (i.e. poor program differentiation).

In the current study, we assess how providers from a national probability sample of schools with middle grades implemented evidence-based school substance use prevention curricula. Based on their reports, we examine implementation of the evidence-based curricula along the five fidelity domains of adherence (including the subdomains of prescribed content and delivery strategies), exposure, quality of delivery, participant responsiveness and program differentiation and we consider all the domains together. We also examine the relationships among the fidelity domains, with the expectation that all domains will be positively related to each other.

Method

Data source

Data are from the second wave of the School-based Substance Use Prevention Programs Study, a longitudinal study of substance use prevention practices in the nation's public schools, with primary focus on the middle school grades [34]. The study was exempted from human subjects review. We selected schools in two phases, the first of which came from a 1997–98 sampling frame from the Quality Educational Database [35]. We defined schools with middle grades as those with a stand-alone sixth

grade, that comprised the fifth and sixth grades only or that included either seventh or eighth grade. Excluded from the frame were schools designated as alternative, charter, vocational/technical or special education, those administered by the U.S. Department of Defense or Bureau of Indian Affairs or those with <20 students. The sampling frame yielded 2273 eligible public schools in the 50 states and District of Columbia. A refreshment sample of 210 public schools using these same inclusion criteria was drawn from a 2002–03 sampling frame maintained by the Common Core of Data [36]. The purpose of this second sampling phase was to maintain the sample's representativeness by accounting for new schools opened in the intervening 5-year time period. Both samples were stratified by population density, school size and poverty level, with equal probabilities of selection within each stratum. Data were collected for the second wave in 2005. School sample characteristics for the current analysis sample are shown in Table I.

Data collection

Prior to data collection, we telephoned each school's administrative staff to identify an appropriate respondent, defined as the most knowledgeable person about substance use prevention in the school who also taught substance use prevention. Most respondents were teachers; others were school counselors, prevention specialists or held other positions. We surveyed these program providers via a secured website after inviting them to participate by a letter that included a prepaid \$10 cash incentive. Those who did not complete the web survey after repeated contacts were mailed a paper copy of the questionnaire; those who did not complete the mailed survey were contacted for a brief telephone interview that contained a reduced set of questions. The overall response rate was 78.2% ($N = 1721$), and the majority (65.2%) responded to the web survey. See Table I for background characteristics of respondents.

We asked providers to identify the substance use prevention curricula they were teaching in the current school year (2004–05) from a list of 27

Table I. Respondent and school sample characteristics (*N* = 342)

Characteristic	% or mean	95% CI
Respondent		
Female	76.65	72.15–81.16
White non-Hispanic	85.23	82.10–88.36
African American non-Hispanic	9.20	6.85–11.54
Other race/ethnicity non-Hispanic	2.76	0.81–4.71
Hispanic	2.61	1.12–4.10
Mean age	44.20 years	43.12–45.28
Mean years teaching substance use prevention	11.50 years	10.60–12.29
School region ^a		
Northeast	18.02	14.32–21.72
Midwest	28.89	24.08–33.69
South	32.31	27.56–37.05
West	20.79	16.49–25.08
Population density of geographic area served ^b		
Urban	21.92	17.39–26.47
Suburban	30.43	25.47–35.39
Rural	47.64	41.55–53.73
School poverty (% of students eligible for free or reduced-price lunch) ^b		
Low (0–14%)	23.09	19.96–26.22
Medium (15–39%)	31.76	27.91–35.61
High (>39%)	45.15	42.02–48.28
School size (number of students in Grades 5–8) ^b		
Small (20–199)	26.11	20.76–31.46
Medium (200–599)	36.94	31.08–42.80
Large (600+)	36.95	31.25–42.64
School race/ethnicity composition ^b		
Majority white	76.87	73.27–80.47
Majority African American	5.52	3.24–7.8
Majority Hispanic	9.29	6.9–11.68
Other majority	2.29	0.28–4.3
No majority	6.03	3.73–8.34

N is unweighted and proportions calculated using weighted data.

^aDefined by US Census regions.

^bDefined based on school data available from the 2004–05 Common Core of Data school file.

universal substance use prevention programs available at that time that targeted middle grade youth. Although not noted as such for respondents, the list included 10 curricula that met criteria for being designated ‘evidence-based’ by any of three national registries of prevention programs. We defined evidence-based curricula as those identified at the time as ‘model’ or ‘effective’ by NREPP [1], as ‘model’ or ‘promising’ by Blueprints for Violence Prevention [37] or as ‘exemplary’ by the Office of Safe and Drug-Free Schools [38]. The curricula were All Stars, keepin’ it REAL, Life-

Skills Training, Lions Quest Skills for Adolescence, Positive Action, Project ALERT, Project Northland, Project Toward No Tobacco Use (TNT), Social Competence Promotion Program for Young Adolescents and Too Good for Drugs. Descriptive information about each program, including journal citations, can be found on NREPP [1]. These programs vary somewhat in the content covered but all share an emphasis on using interactive delivery strategies, such as demonstration and practice of skills and role plays, in contrast to didactic methods of instruction [13, 15].

Because of prior evidence that many schools administered two or more substance use prevention programs [34], providers were asked to select from the list all curricula they were currently teaching, and in a subsequent question, they were asked to identify the one curriculum they were teaching the most. Providers then were directed to modules of questions pertaining to how they taught the curriculum. For three curricula, All Stars, Life Skills Training and Project ALERT, the modules incorporated the curriculum name into the questions and included other curriculum-specific detail as appropriate (e.g. specific lesson names). For all other curricula, respondents answered parallel questions where the referent was 'the curriculum you are using the most with students in middle or junior high grades'.

Analysis sample

We restricted the analysis sample to providers who reported teaching 1 of the 10 universal evidence-based substance use prevention curricula the most in the 2004–05 school year ($N = 399$). Because some questions used to form the measures were not included in the abbreviated telephone interview, we further restricted the sample to those who completed the survey by Web or mail ($N = 342$; 85.7% of the eligible sample).

Fidelity of implementation measures

We formed measures of program adherence (from a combination of two separately constructed measures of content and delivery strategies), exposure, quality of delivery, participant responsiveness and program differentiation from providers' responses to questions about how they implemented their substance use prevention curriculum. Implementation adherence, exposure and quality of delivery were assessed with sets of variables that were combined to form summary measures of the domains. Participant responsiveness and program differentiation were assessed by one measure each. For each of the five domains, as well as for the two adherence subdomains of content and delivery strategies, we created a dichotomous measure that contrasted those who demonstrated fidelity on the domain with

those who did not. A description of the fidelity measures and variables, including the cut points for operationalizing fidelity, is provided in Table II.

As noted on the table, the two measures of the content subdomain of adherence and of exposure were tailored to features of the specific curricula. We provide additional detail about these two measures here. For details about each curriculum needed to construct the measures, we obtained descriptions of the curricula from NREPP, program Web sites, program manuals and in some cases from personal communication with program developers.

The possible content areas targeted by curricula were classified as information (e.g. drug use consequences, social and media influences), substance use refusal skills, personal competency skills (e.g. decision making) and positive affect and beliefs (e.g. improving self-esteem, reinforcing positive attitudes). The measure of the content subdomain of adherence was coded dichotomously to contrast those providers who were covering all content areas emphasized in the focal curriculum at relatively high levels (i.e. covered each content area on average in 'some' lessons or more) with those who were covering the content areas at relatively low levels (i.e. covered each content area on average in fewer than 'some' lessons). For LifeSkills Training, Lions Quest Skills for Adolescence, Project ALERT, Project Northland, Project TNT and Too Good for Drugs, the content areas emphasized were information, refusal skills and personal competency skills. For All Stars and Positive Action, the content areas emphasized were personal competency skills and positive affect and beliefs.

Program exposure was measured by a composite of the number and frequency of lessons taught. All providers answered a single question with response options for the exact number of lessons taught, up to '≥16'. Only two curricula, Lions Quest Skills for Adolescence and Positive Action, included >16 lessons, and those who reported teaching at least this many lessons were coded as teaching all of them. Those using All Stars, Life Skill Training and Project ALERT also were asked how much of each lesson they had taught using a checklist of all lessons, which included the specific lesson name and

Table II. *Fidelity measures*

Fidelity measure and variables	Constituent variables or number of items	Response categories
Adherence	Composite of (a) content and (b) interactive delivery strategies	1 = implemented prescribed content and interactive delivery strategies, 0 = not
Content ^a	Composite of frequency of content areas	1 = covered emphasized content areas in 'some' or 'most' lessons (i.e. average frequency ≥ 3), 0 = not
Frequency of information content	5	1 = never to 4 = most lessons
Frequency of refusal skills content	2	1 = never to 4 = most lessons
Frequency of personal and social competency skills content	2	1 = never to 4 = most lessons
Frequency of positive affect and beliefs content	3	1 = never to 4 = most lessons
Interactive delivery strategies	Composite of (a) frequency of interactive and (b) frequency of non-interactive strategies	1 = used interactive strategies in 'most lessons' (i.e. frequency = 4) and more than non-interactive strategies (i.e. frequency < 4), 0 = not
Frequency of interactive strategies	4	1 = never to 4 = most lessons
Frequency of non-interactive strategies	3	1 = never to 4 = most lessons
Exposure ^a	Composite of (a) number lessons taught and (b) frequency of lessons	1 = taught all lessons at recommended frequency, 0 = not
Number of lessons taught	1	1 = none to 17 = 16 or more
Frequency of lessons	1	1 = 1 lesson per month or less often to 5 = daily
Quality of program delivery	Composite of (a) teacher encouragement of students and (b) teacher confidence	1 = 'usually' or 'always' encourages students (i.e. ≥ 4) and 'agrees' or 'strongly agrees' (i.e. ≥ 4) is confident teaching the program, 0 = not
Teacher encouragement of students	2	1 = never to 5 = always
Teacher confidence	2	1 = strongly disagree to 5 = strongly agree
Participant responsiveness	2	1 = agrees or strongly agrees students responded enthusiastically (i.e. ≥ 4), 0 = not
Program differentiation	28	1 = used only one evidence-based program, 0 = not

^aTailored to specifications of each evidence-based program.

a brief description of each. Because of the presumed greater validity of an exposure measure based on the list of specific lessons compared with the general measure of the number of lessons taught and because almost three-quarters of the sample used one of these three curricula, we used the lesson list when available to code the number of sessions taught (regardless of how much of each session was taught). Comparison of the two exposure measures showed that providers reported implementing more lessons based on the specific lists than on the general question.

Analysis

We report the proportion of school providers using each evidence-based substance use prevention curriculum. We provide descriptive statistics for the fidelity measures averaged across providers (and thus curricula) and report the percent of providers achieving each specific fidelity domain and all five domains considered in aggregate. For all estimates, we provide 95% confidence intervals (CIs). We assessed the relationships between pairs of fidelity domains using Rao–Scott chi-square tests. Because

of missing data on some items, the sample sizes for the fidelity measures and variables ranged from 307 to 342. Non-response and post-stratification adjustments were used to adjust slight discrepancies between populations and samples in the full file of 1721 cases. All analyses were based upon these weighted data. The weights had a negligible effect on variance/standard errors. All analyses were conducted using the SurveyFreq and SurveyMeans procedures of SAS 9.1.3 [39].

Results

School providers reported using 8 of the 10 evidence-based curricula (Table III), with Project ALERT and LifeSkills Training by far the most common choices. Substantial variability in fidelity of implementation was present across the five fidelity domains, as well as in the two subdomains of adherence, namely content and delivery strategies (Table IV). Just more than one-quarter of providers demonstrated fidelity of implementation on the composite measure of adherence, although a substantially higher percentage reported fidelity on the constituent dimension of implementing the prescribed content compared with the dimension concerning the use of prescribed delivery strategies. The latter tapped the frequent use of interactive strategies at higher levels than non-interactive strategies. The two dimensions of adherence were significantly related to each other [χ^2 (1 d.f.) = 10.81, $P < 0.001$], such that providers who frequently used interactive teaching methods were more likely to implement the prescribed content.

Only about one-third of providers achieved fidelity on the exposure domain, meaning they implemented all the curriculum lessons on the recommended schedule. Even fewer providers reported implementing only the focal evidence-based curriculum during the same school year (program differentiation). In contrast, large percentages of providers reported high levels of engagement in teaching the curricula (quality of delivery) and high participant responsiveness. Almost no providers were coded as fully demonstrating fidelity on all five domains considered together.

Table III. Evidence-based substance use prevention curricula usage in 2004–05 school year among providers using an evidenced-based curriculum as their primary curriculum ($N = 342$)

Evidence-based curriculum	<i>N</i>	%	95% CI
All Stars	9	3.0	1.6–4.4
LifeSkills Training	117	36.3	31.0–41.7
Lions Quest Skills for Adolescence	27	6.9	4.2–9.7
Positive Action	6	1.7	0.21–3.1
Project ALERT	136	38.9	33.4–44.4
Project Northland	11	2.7	1.0–4.3
Project TNT	4	1.2	0.0–2.4
Too Good for Drugs	32	9.3	6.2–12.3

No providers reported using keepin' it REAL or Social Competence Promotion Program for Young Adolescents as their primary program. *N*s are unweighted and proportions calculated using weighed data.

Relationships among the fidelity dimensions showed that teachers who reported high adherence were significantly more likely to report high-quality delivery [χ^2 (1 d.f.) = 13.44, $P < 0.001$] and high student responsiveness [χ^2 (1 d.f.) = 15.93, $P < 0.0001$]. High-quality delivery was significantly associated with full curriculum exposure [χ^2 (1 d.f.) = 4.39, $P < 0.05$] and high student responsiveness [χ^2 (1 d.f.) = 79.21, $P < 0.0001$], but inversely associated with implementing only the focal curriculum [χ^2 (1 d.f.) = 3.96, $P < 0.05$]. Other relationships among fidelity domains were not statistically significant.

Post hoc analyses

We conducted two sets of *post hoc* analyses to probe findings related to exposure and program differentiation. We created alternative measures of exposure for Project ALERT, LifeSkills Training and All Stars' providers using the additional detailed information obtained from the curriculum-specific lesson lists. According to these lists, providers implemented an average of 85.6% (95% CI = 83.1–90.0%) of the lessons. They also reported teaching on average 'most' of the lesson materials [range = 1 (none) to 4 (all); mean = 2.9, 95% CI = 2.6–3.2] for each lesson. These figures suggest higher exposure than indicated by our original

Table IV. Descriptive statistics for provider-reported measures of fidelity of implementation

Fidelity measure and variables	Mean or %	95% CI
Adherence, % implemented prescribed content and interactive delivery strategies	27.7	22.6–32.8
Content, %	70.8	65.8–75.9
Frequency of information content ^a	3.1	3.1–3.2
Frequency of refusal skills content ^a	2.8	2.7–2.9
Frequency of personal and social competency skills content ^a	3.3	3.2–3.3
Frequency of positive affect and beliefs content ^a	3.1	3.0–3.1
Interactive delivery strategies, %	33.4	28.1–38.7
Frequency of interactive strategies ^a	3.0	2.9–3.1
Frequency of non-interactive strategies ^a	3.1	3.0–3.1
Exposure, % fully implemented curriculum	35.8	30.5–41.2
Taught all lessons, %	55.4	50.1–60.7
Taught lessons on recommended schedule, %	56.9	51.4–62.5
Quality of program delivery, % encouraged students and confident delivering	85.3	81.5–89.2
Provider encouragement of students ^b	4.5	4.3–4.5
Provider confidence ^b	4.0	3.9–4.1
Participant responsiveness, %	80.3	75.7–85.0
Program differentiation, % used only the focal curriculum	15.3	11.5–19.1
Fidelity, % demonstrating adherence, exposure, quality of program delivery, participant responsiveness and program differentiation criteria	1.3	0.0–2.7

Sample sizes range from $N = 307$ to 342 because of missing data on items. Sample includes school providers who taught an evidence-based substance use prevention curriculum 'the most' in the 2004–05 school year. All estimates use weighted data.

^aRange = 1–4.

^bRange = 1–5.

measure that was operationalized for the entire sample and which took into account each curriculum's suggested implementation schedule. The alternative

findings are noteworthy because the majority of providers (78.2% of the sample) used one of these three curricula.

To probe the program differentiation finding, we found that 33.9% of the providers (95% CI = 28.8–39.0%) reported using one or more additional evidence-based curricula, 58.5% (95% CI = 53.2–63.9%) taught one or more curricula not designated as evidence based and 47.8% (95% CI = 42.3–52.4%) used a locally developed program or set of materials. Providers could have reported using any one or more of these other programs. Considered together, providers were more likely to supplement their focal evidenced-based curriculum with non-evidence based rather than evidence-based programs.

Discussion

Adherence and exposure constitute the two domains of implementation fidelity at the heart of whether a program is implemented as intended by its developers. Yet far fewer providers of evidenced-based substance use prevention curricula achieved fidelity on these domains relative to the proportions who achieved fidelity on quality of delivery or participant responsiveness; providers were least likely to achieve fidelity on the program differentiation domain. Only about one-third of providers delivered the full curriculum on the recommended schedule and only one-quarter were found to adhere to both the prescribed content and delivery strategies. The percent of providers rated as adherent was driven by the subdomain of delivery strategies: only about one-third of providers delivered interactive strategies at the prescribed frequency. This estimate is considerably greater, however, than the 17% who used interactive delivery strategies that we reported in the initial round of the study conducted 6 years earlier [14]. While sample differences somewhat compromise the comparison, the findings suggest both progress in the uptake of interactive delivery strategies and the challenges that remain to school providers in using these methods.

As with adherence, there is some opportunity to take encouragement from our findings about exposure. While only around one-third of providers in the full sample reported implementing the whole curriculum on the schedule suggested by program developers, the percentages achieving high exposure were greater when using an alternative measure operationalized for the large subsample of providers using Project ALERT, LifeSkills Training and All Stars. These providers completed an average of 86% of program lessons, which compares well with exposure or dosage estimates from evaluation research [22, 26, 27].

The lower percentages of respondents achieving fidelity on the domains of adherence and exposure compared with the percentages on quality of delivery and participant responsiveness are not particularly surprising in that the former represent assessments of program implementation actions, whereas the latter represent more global assessments of performance. Of perhaps greater significance than the modest levels of adherence and exposure is the finding that these two domains were unrelated. The lack of association likely reflects the findings noted above that suggest that providers deliver curriculum lessons but not necessarily while following the prescribed delivery strategies. Notably, however, both adherence and exposure were significantly associated with quality of delivery. Providers who reported higher quality delivery—in that they were more confident of their ability to teach their evidence-based curriculum and were more encouraging of their students' participation—were more likely to report adhering to prescribed content and delivery strategies as well as to implement the full curriculum. These providers also were more likely to report that their students actively participated in the curriculum. Provider engagement may be central to program fidelity.

Unexpectedly, those providers who were high versus low on quality of delivery of the focal evidence-based curriculum were more likely to deliver other substance use prevention programs in the same school year. Given that these providers were more likely to be adherent and engaged in teaching substance use prevention, perhaps they intended to

enhance the learning experience for students with supplementary materials. Indeed, Rogers [40] noted that 're-invention', whereby an intervention is modified when implemented, is common and may not be counterproductive when the adaptations are intentionally meant to address local needs and do not impair the underlying theoretical model. However, the tendency of these providers to use non-evidence-based curricula and locally developed materials more often than other evidence-based curricula sounds a cautionary note.

Measurement issues provide a caveat to any conclusions from our findings. As already discussed, our data yielded different conclusions about exposure fidelity depending on the measure we used. As another example, following definitions used in Tobler's meta-analyses of school drug prevention programs, we included class discussions as an indicator of non-interactive methods because these discussions tend to involve communication between teachers and students rather than discussion among peers [13, 15]. Teachers reporting the use of class discussions could be grouping teacher-led and peer-focused discussions. Had we included class discussion as an indicator of interactive strategies, our estimates of adherence would have been higher. These examples illustrate that the strategy used to operationalize fidelity measures will inevitably lead to varying estimates of fidelity. They also point to problems related to the lack of standard definitions of fidelity in this emerging field of enquiry.

An additional measurement consideration relates to the source of information. Observational data are less subject to social desirability bias and thus may provide more valid estimates of fidelity than the self-reported data used here [23–25]. Our estimates, therefore, may be inflated. On the other hand, our participants were not involved in research to evaluate any particular program and thus may have felt less incentive to respond favorably. Furthermore, providers may have been simply unaware of the nature and extent to which their administration of evidence-based curricula differed from prescribed guidelines and thus less likely to inflate their responses.

Another measurement concern is the effect on recall of how recently providers taught their

curricula. With data collection in the Spring of the 2004–05 school year, many providers likely implemented their curriculum during the Fall. Their recollection of how many lessons they implemented may thus have been compromised. While both observational data and implementation checklists collected immediately from providers would have improved our assessment of fidelity, these methods were not practical given a national sample, conditions of real-world implementation and the number of evidence-based curricula in use.

Our findings shed light on fidelity of implementation of evidence-based school substance use prevention curricula as experienced by providers working under natural conditions. With fidelity of implementation under research conditions as the standard referent, it would be unreasonable to expect providers to achieve complete fidelity on all domains, which has not been demonstrated even under the most rigorous research conditions [6]. Yet, reasonably high expectations are appropriate and necessary if curricula are to have their intended effects on youth substance use. Our results suggest that until higher levels of adherence to content and delivery strategies can be achieved, expectations must be tempered. The findings also suggest the need for continued emphasis on fidelity in program materials, training and on-going technical support with particular attention to supporting use of the interactive delivery methods called for by the programs' developers. Perhaps most importantly, we need research that examines why providers do not deliver curricula as intended to inform both curriculum development and training for existing programs.

Funding

National Institute on Drug Abuse (NIDA R01 DA016669 to C.L.R.).

Conflict of interest statement

None declared.

References

1. Substance Abuse and Mental Health Services Administration. *NREPP: SAMSHA's National Registry of Evidence-based Programs and Practice*. Available at: <http://www.nrepp.samhsa.gov/>. Accessed 30 June 2010.
2. *No Child Left Behind Act of 2001*. 2002; Pub. L. No. 107-110, 115 Stat 1425.
3. U.S. Department of Education. Safe and drug-free schools program. Notice of final principles of effectiveness. *Fed Regist* 1998; **63**: 29902–6.
4. Ringwalt C, Vincus AA, Hanley S *et al*. The prevalence of evidence-based drug use prevention curricula in U.S. middle schools in 2008. *Prev Sci* 2011; **12**: 63–69.
5. Botvin GJ. Advancing prevention science and practice: challenges, critical issues, and future directions. *Prev Sci* 2004; **5**: 69–72.
6. Durlak JA, DuPre EP. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am J Community Psychol* 2008; **41**: 327–50.
7. Dusenbury L, Brannigan R, Falco M *et al*. A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Educ Res* 2003; **18**: 237–56.
8. Elliott DS, Mihalic S. Issues in disseminating and replicating effective prevention programs. *Prev Sci* 2004; **5**: 47–53.
9. Rohrbach LA, Grana R, Sussman S *et al*. Type II translation: transporting prevention interventions from research to real-world settings. *Eval Health Prof* 2006; **29**: 302–33.
10. Botvin GJ, Baker E, Dusenbury L *et al*. Long term follow up results of a randomized drug abuse prevention trial in a white middle class population. *J Am Med Assoc* 1995; **273**: 1106–12.
11. Dane AV, Schneider BH. Program integrity in primary and secondary prevention: are implementation effects out of control? *Clin Psychol Rev* 1998; **18**: 23–45.
12. Stead M, Stradling R, MacNeil M *et al*. Implementation evaluation of the blueprint multi-component drug prevention programme: fidelity of school component delivery. *Drug Alcohol Rev* 2007; **26**: 653–64.
13. Tobler NS, Stratton HH. Effectiveness of school-based drug prevention programs: a meta-analysis of the research. *J Prim Prev* 1997; **18**: 71–128.
14. Ennett ST, Ringwalt CL, Thorne J *et al*. A comparison of current practice in school-based substance use prevention programs with meta-analysis findings. *Prev Sci* 2003; **4**: 1–14.
15. Tobler NS, Roona MR, Ochshorn P *et al*. School-based adolescent drug prevention programs: 1998 meta-analysis. *J Prim Prev* 2000; **20**: 275–336.
16. Rohrbach LA, Graham JW, Hansen WB. Diffusion of a school-based substance abuse prevention program: predictors of program implementation. *Prev Med* 1993; **22**: 237–60.
17. Orwin RG. Assessing program fidelity in substance abuse health services research. *Addiction* 2000; **95**: S309–27.
18. Steckler AB, Linnan L. *Process Evaluation for Public Health Interventions and Research*, 1st edn. San Francisco: Jossey-Bass, 2002.

19. Skara S, Rohrbach LA, Sun P *et al.* An evaluation of the fidelity of implementation of a school-based drug abuse prevention program: Project toward No Drug Abuse (TND). *J Drug Educ* 2005; **35**: 305–29.
20. Hansen WB, Graham JW, Wolkenstein BH *et al.* Program integrity as a moderator of prevention program effectiveness: results for fifth grade students in the adolescent alcohol prevention trial. *J Stud Alcohol* 1991; **52**: 568–79.
21. Tortu S, Botvin GJ. School-based smoking prevention: the teacher training process. *Prev Med* 1989; **18**: 280–9.
22. Mihalic SF, Fagan AA, Argamaso S. Implementing the Life-Skills Training drug prevention program: factors related to implementation fidelity. *Implement Sci* 2008; **3**: 5.
23. Lillehoj CJ, Griffin KW, Spoth R. Program provider and observer ratings of school-based preventive intervention implementation: agreement and relation to youth outcomes. *Health Educ Behav* 2004; **31**: 242–57.
24. Moncher FJ, Prinz RJ. Treatment fidelity in outcome studies. *Clin Psychol Rev* 1991; **11**: 247–66.
25. Resnicow K, Smith M, Davis M. How best to measure implementation of health curricula? A comparison of three measures. *Health Educ Res* 1998; **13**: 239–50.
26. Pentz MA, Trebow EA, Hansen WB *et al.* Effects of program implementation on adolescent drug use behavior: the midwestern prevention project (MPP). *Eval Rev* 1990; **14**: 264–89.
27. Rohrbach LA, Dent CW, Skara S *et al.* Fidelity of implementation in Project towards No Drug Abuse (TND): a comparison of classroom teachers and program specialists. *Prev Sci* 2007; **8**: 125–32.
28. Sloboda Z, Stephens P, Pyakuryal A *et al.* Implementation fidelity: the experience of the adolescent substance abuse prevention study. *Health Educ Res* 2009; **24**: 394–406.
29. Hansen WB. Pilot test results comparing the all star program with seventh grade D.A.R.E.: program integrity and mediating variable analysis. *Subst Use Misuse* 1996; **31**: 1359–77.
30. Harrington NG, Giles SM, Hoyle RH *et al.* Evaluation of the all stars character education and problem behavior prevention program: effects on mediator and outcome variables for middle school students. *Health Educ Res* 2001; **28**: 533–46.
31. Dusenbury L, Brannigan R, Hansen WB *et al.* Quality of implementation: developing measures crucial to understanding the diffusion of preventive interventions. *Health Educ Res* 2005; **20**: 308–13.
32. Hallfors D, Godette D. Will the ‘principles of effectiveness’ improve prevention practice? early findings from a diffusion study. *Health Educ Res* 2002; **17**: 461–70.
33. Hansen WB, McNeal RB. Drug education practice: results of an observational study. *Health Educ Res* 1999; **14**: 85–97.
34. Ringwalt CL, Ennett S, Vincus A, *et al.* The prevalence of effective substance use prevention curricula in U.S. middle schools. *Prev Sci* 2002; **3**: 257–65.
35. Quality Education Data Inc. *QED National Education Database: Data Users Guide, Version 4.6*. Denver, CO: Author, 1998.
36. National Center for Education Statistics. *Public Elementary/Secondary School Universe Survey Data, 2002–03*. [data file]. Available at: <http://nces.ed.gov/ccd/pubagency.asp>. Accessed: 21 February 2007.
37. Center for the Study and Prevention of Violence. *Blueprints for Violence Prevention: Model Programs and Promising Programs*. Available at: <http://www.colorado.edu/cspv/blueprints/modelprograms.html> and <http://www.colorado.edu/cspv/blueprints/promisingprograms.html>. Accessed February 2007.
38. Safe, Disciplined, and Drug-Free Schools Expert Panel. *Exemplary Programs*. Available at: http://www.ed.gov/offices/OERI/ORAD/KAD/expert_panel/2001exemplary_sddfs.html. Accessed February 2007.
39. SAS [computer program]. *Version 9.1.3*. Cary, NC: SAS Institute, Inc, 2003.
40. Rogers EM. *Diffusions of Innovations*, 4th edn. New York, NY: The Free Press, 1995.