Evaluating school capacity to implement new programs

Cynthia Roberts-Graya,*, Phyllis M. Gingissb, Melynda Boermb

aThird Coast R&D Inc., 2200 Market Ste 810, Galveston, TX 77550, USA
bHealth Network for Evaluation and Training Systems (HNETS), Department of Health and Human Performance, University of Houston, Houston, TX, USA

Received 3 February 2007; received in revised form 4 April 2007; accepted 8 April 2007

Abstract

An eight-factor survey-based Bayesian model (Bridge-It) for assessing school capacity to implement health and education programs was tested in secondary analyses of data from 47 schools in the Texas Tobacco Prevention Initiative (TTPI). Bridge-It was used during the pre-implementation phase and again at mid-course of the TTPI 2 years later. Achieved implementation status was evaluated in follow-up almost 4 years after the start of the TTPI. The Bridge-It score aggregated across all eight of the capacity factors predicted both quality of adherence to the Guidelines for School Programs to Prevent Tobacco Use and Addiction and quantity of implementing activity. The school-based leadership factor was an independent predictor of quality of adherence whereas the facilitation processes factor predicted quantity of implementing activity. Integration of Bridge-It, or comparable multi-attribute tools, into the planning and evaluation of school-centered programs can increase understanding of factors that influence implementation and provide guidance for capacity building.

1. Introduction

Implementation is the “bridge” between a school program and its impact on students and their families (Berman & McLaughlin, 1976; also cited in Dusenbury, Brannigan, Falco, & Hansen, 2003). No matter how effective it proved in the research lab or field trial, a school-centered health or education program cannot produce its intended benefits until it is effectively integrated into actual procedures and practices at the campus level.

Effective implementation, however, is elusive. Too often, promising programs fall into an implementation gap, adopted but not used or only partially integrated into actual practice. Inspection of the research literature suggests as many as half of schools into which innovative and evidence-based programs are introduced fail to implement with sufficient scope or fidelity to expect the promised benefits will accrue to students, families, and communities. This challenge becomes more acute when evidence-based practices and guidelines include program components that span multiple aspects of school functioning, such as policy change and enforcement, classroom instruction, counseling support, and family involvement. For examples, see Berman & McLaughlin, 1978; Busto, Wight, Hart, & Scott, 2002; Elliott & Mihalic, 2004; Eslea & Smith, 1998; Gingiss, Roberts-Gray, & Boerm, 2006; Gottfredson & Gottfredson, 2002; Hahn, Noland, Rayens, & Christie, 2002; Hallfors & Godette, 2002; Kumpfer, 2002; Pentz, 2003; Ringwalt et al., 2003; Taggart, Bush, Zuckerman, & Theiss, 1990).

Part of the explanation for failed or low levels of implementation is lack of capacity at the campus level to implement the new program or technology (Fullan, 2005; Johnson, Hays, Center, & Daley, 2004; Kaffarian, Robinson, Compton, Davis, & Volkow, 2004). Elements of school capacity to implement new programs and policies...
include support by the principal; the teachers’ sense of efficacy; implementers’ ability to communicate with program participants; staff morale; the general school culture; quality of leadership; availability of time, money, or other resources; presence or amount of turbulence in the implementing environment; and other organizational characteristics (Bosworth, Gingiss, Pothoff, & Roberts-Gray, 1999; Dusenbury et al., 2003; Kallestad & Olweus, 2003). Program characteristics such as complexity and relative advantage; facilitation processes such as adaptive planning and implementer training; and factors external to the school such as state- and district-level policies and mandates are other elements to be considered in understanding and building school capacity to implement a new health or education program (Berman & McLaughlin, 1976; Blake et al., 2005; Boerm, Gingiss, & Roberts-Gray, 2007; Bosworth et al., 1999; Greenberg & Walls, 2003; Han & Weiss, 2005; Payne, Gottfredson, & Gottfredson, 2006; Ringwall et al., 2003).

To enable and encourage successful implementation of school programs for prevention of tobacco use and addiction, evaluation and planning for the Texas Tobacco Prevention Initiative (TTPI) included a focus on schools’ capacity to implement best practices identified in the Guidelines for School Health Programs to Prevent Tobacco Use and Addiction (school guidelines, CDC, 1994, 2004). Tobacco use is one of the six health behaviors that contribute most to the leading causes of mortality in the United States (Kann, Brener, & Allensworth, 2001). Frequently these behaviors are established during youth. Nationwide, 22.3% of high school students and 8.1% of middle school students are current cigarette smokers (CDC, 2005). School health education programs play an important role in reducing adolescent tobacco use by increasing student knowledge, positive attitudes and peer resistance skills, and therefore lowering the levels of youth smoking (Engquist et al., 1994). When properly implemented, school programs can lower smoking prevalence by 25–60% (Meshack et al., 2004; National Cancer Policy Board, 2000; Rhode et al., 2001). The school guidelines were designed as a set of recommendations for ensuring a quality school program to prevent tobacco use. These recommendations constitute a “bundle” of program components which necessitate school adaptation and accommodation on multiple dimensions and at multiple levels.

To build schools’ capacity to implement programs consistent with the school guidelines, the TTPI, which is administered through the Texas Department of State Health Services (TDSHS) using funds from the Texas Tobacco Settlement, awarded small ($2000) competitive grants to reimburse program expenses and also provided initial training, guidance, and materials to selected schools in East Texas (TDSHS, 2001). Evaluation of the TTPI included studies to monitor schools’ capacity for achieving and sustaining successful implementation of tobacco prevention and control programs and studies to track implementation status (Gingiss, Boerm et al., 2006; Gingiss, Roberts-Gray et al., 2006; Boerm et al., 2007).

The Bridge-It system (Bosworth et al., 1999) was adopted to assess capacity to implement school programs in the TTPI. Bridge-It originally was developed using an expert panel and an empirical group process technique to integrate the wide range of influential variables into a model to help schools plan for and monitor key elements of the implementation process for school programs. The system includes an 8-factor, 36-item survey to analyze capacity for program implementation and a companion Bayesian model that uses the survey data to estimate the likelihood of implementation success.

The current study uses results from TTPI evaluation studies conducted at three points in time. The baseline or pre-implementation study was conducted to document the then current amount of implementing activity for components of the school guidelines and assess schools’ capacity to achieve and maintain proper implementation of the recommended best practices. Two years later, mid-course assessments were conducted to assess capacity for continued implementation success. Almost 4 years after the pre-implementation assessment, a more comprehensive follow-up assessment of implementation status was conducted using the survey tools adapted from the School Health Education Profile—Tobacco Module (SHEP-TM, Centers for Disease Control and Prevention, 2001). Data from the sequence of three evaluation studies were submitted to secondary analyses to examine the utility and predictive validity of the Bridge-It system. Our hypotheses were that (1) the Bridge-It model for assessing school capacity to implement new health and education programs provides valid predictors of implementation status measured nearly 4 years after program start-up; and (2) Bridge-It’s multi-factor approach for measuring capacity has utility because different factors are predictive of qualitative versus quantitative aspects of achieved implementation status.

2. Method

2.1. Participants

Schools were selected for the current study if they were (a) among the initial 111 schools that began their participation in the TTPI during fall 2000 or spring 2001; (b) participated in the baseline capacity assessment conducted August through October 2000 and/or mid-course capacity assessment in August 2002; and (c) participated in the survey of implementation status in April 2004. Forty-seven schools met these eligibility criteria. The survey forms for capacity assessments were distributed to the designated campus-level coordinator for the TTPI contract. These individuals described their position on the campus as teachers (51%), counselors (17%), school nurses (13%), administrators (8%), safe and drug-free coordinators (6%), or “other” such as social
worker (5%). The survey forms for implementation status assessment in 2004 were distributed to the school principals and school health coordinators.

Concerned about possible bias introduced because more than half of the initial schools did not respond to the survey questionnaires at one or more of the data collection points, we conducted analysis of variance to compare enrollment and characteristics of the student body of schools that met eligibility criteria for the current study and those that did not. We found no significant differences. Percent economically disadvantaged, for example, was $48.66 \pm 30.17$ at schools that did and $46.96 \pm 21.77$ at those that did not participate in all three of the evaluation studies. The schools represented a diversity of campuses and academies serving students in middle schools, intermediate schools, junior high schools, high schools, and alternative education programs distributed across school districts in urban, suburban, and rural areas of East Texas. Demographic characteristics of the 47 schools in the current study are displayed in Table 1.

As a further check on potential self-selection bias, we compared the implementation status scores obtained in 2004 follow-up for schools that participated only in the follow-up ($n = 13$), in the follow-up and only one of the capacity assessments ($n = 19$), or all three of the evaluation studies ($n = 28$). We also compared the overall pre-implementation capacity scores for schools that participated in the implementation status evaluation in 2004 ($n = 35$) and those that participated in the pre-implementation capacity assessment in 2000 but not in the implementation status evaluation in 2004 ($n = 56$). Again, we found no significant differences. The average implementation quality score in 2004, for example, was $62.94 \pm 14.29$, $73.26 \pm 11.25$, and $71.30 \pm 12.86$, respectively, for schools that participated only in the follow-up, those in the follow-up and one of the capacity assessments, and those that participated in all three of the evaluation studies. The average pre-implementation capacity score in 2000 was $19.51 \pm 24.43$ and $19.85 \pm 26.02$, respectively, for schools that did and those that did not participate in the implementation status evaluation in 2004.

Because the sample of schools with pre-implementation, mid-course, and follow-up data was fewer than 30 and because there was minimal change in capacity scores from pre-implementation to mid-course, we combined the datasets for pre-implementation and midcourse capacity assessments. For each of the 47 schools in the current study, we used the oldest available assessment to describe capacity to implement and to predict implementation status. To investigate trends the analyses were repeated using only the data for the 28 schools that participated in all three of the evaluation studies.

2.2. The Bridge-It measures of capacity to implement

The eight factors queried in the Bridge-It capacity survey are multi-attribute constructs with each attribute operationally defined as a single item with its own self-anchored scale. The eight factors include three factors measuring aspects of campus-level infrastructure—resources, school-based leadership, and implementer characteristics; two factors addressed to aspects of the program—compatibility and innovation characteristics; two factors to assess forces external to the campus—external leadership and external environment; and one implementation process factor labeled facilitation processes.

Survey items to assess resources ask about adequacy of (a) access to needed materials and supplies, (b) staffing, (c) funding, (d) daily time allocated for planning and follow-through, (e) number of years allowed for planning and trying out the program, and (f) facilities.

School-based leadership items ask (a) how actively engaged the principal is with the program, (b) whether a person has been identified to provide program leadership at the campus and been given time for the task, (c) nature of the working relationship between the principal and the campus-level program leader, and (d) how important the program is to the principal in terms of priority.

Items to measure implementer characteristics ask about (a) level of professional preparation of staff for their role in implementing the program, (b) breadth of commitment to student health, (c) breadth of skills and experience in the specific techniques used in the program, (d) willingness to try the program; (e) compatibility of the program’s demands with other job expectations and requirements, (f) implementers’ perceptions of compatibility with their professional identity, and (g) implementers’ beliefs about whether the program is better than what they were doing before it was adopted.

Compatibility is measured with questions about the extent to which the program is consistent with the school’s (a) priorities, (b) structure, (c) student needs, (d) culture, and (e) how successful the school has been in implementing other innovations.

Innovation characteristics is assessed with questions asking (a) how complex the program is, (b) how easy it

| Table 1 Demographic characteristics of the schools ($n = 47$) |
|-----------------|-----------------|----------------|
| Total number of students | Mean (Std. dev.) | Range |
| Ethnicity of the student body | 1227 (776) | 56–3072 |
| % African American | 24 (22) | 0–94 |
| % Hispanic | 33 (27) | 0–100 |
| % White | 37 (29) | 0–89 |
| % Asian | 5 (8) | 0–38 |
| Other characteristics of the student body | | |
| % economic disadvantaged | 48 (28) | 2–100 |
| % at risk for school failure | 45 (23) | 4–100 |

aData source: Texas Education Agency Public Education Information Management System (PEIMS) http://www.tea.state.tx.us, accessed 17 December 2005
will be for people to implement the program, and (c) how much of an improvement the program will be over what is currently being done.

The single survey item addressing external leadership asks about level of engagement of a person at district level to facilitate, coordinate, support, and advocate for the program.

External environment is measured with questions about (a) level of disruption or turmoil outside the school, (b) level of support from parents and community, (c) level of community opposition, (d) consistency of the program with federal, state, and district policies and mandates, and (e) how difficult it is to negotiate bureaucratic hurdles to get the program implemented.

Survey items to assess facilitation processes ask about (a) the presence of a written plan, (b) provision for ongoing training for staff and faculty, (c) the level of technical assistance and coaching available, (d) availability and periodicity of coaching and technical assistance, and (e) how many communication channels exist for regular two-way communication among persons involved with implementing the innovation.

The expert panel that participated in the group process to develop Bridge-It defined the questions, the response options, and the scoring procedures (Bosworth et al., 1999). Response options for each survey question are assigned ordinal codes from 1 = least desirable response to 5 = most desirable response. “Don’t know” or “No answer at this time” are coded 0. Although the initial coding was developed to represent an ordinal ranking, the scoring procedure for the factors treats the item codes as categorical. It was the consensus of the expert panel that responses with codes of 4 or 5 would be positively associated with future implementation success.

Factor scores are calculated by counting the number of items with codes of 4 or 5 for that factor’s constituent items and comparing against the criterion values established by the expert panel for scores of “high,” “medium,” or “low.” The compatibility factor, for example, has five items and is scored “high” when four or five, “medium” when two or three, and “low” when one or no items obtain codes of 4 or 5. This scoring procedure assigns the code 0 for “don’t know” or “no response at this time” to the same category as responses with codes of 1 or 2 or 3. Comparison of the proportion of 0 codes in the capacity assessments for schools with the highest \( n = 20 \) and lowest \( n = 20 \) implementation quality scores in 2004 provided empirical support for the appropriateness of this procedure for scoring. Schools with the lowest implementation quality scores in 2004 had significantly higher proportion of 0 codes for capacity assessment items—35% versus 16%, \( \chi^2 \) 8.53, \( p < 0.01 \). Responses of “don’t know” or “no response at this time” to questions assessing school capacity to implement the new program were associated with lesser quality of implementation of the program at follow-up 2 and 4 years later (see also Gingiss & Roberts-Gray, 2002; Roberts-Gray & Gingiss, 2002).

<table>
<thead>
<tr>
<th></th>
<th>Low (%)</th>
<th>Med (%)</th>
<th>High (%)</th>
<th>Mean</th>
<th>Std. dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compatibility</td>
<td>32</td>
<td>21</td>
<td>47</td>
<td>2.15</td>
<td>0.88</td>
</tr>
<tr>
<td>External leadership</td>
<td>43</td>
<td>32</td>
<td>25</td>
<td>1.83</td>
<td>0.82</td>
</tr>
<tr>
<td>Innovation</td>
<td>43</td>
<td>38</td>
<td>19</td>
<td>1.77</td>
<td>0.76</td>
</tr>
<tr>
<td>School-based leadership</td>
<td>46</td>
<td>43</td>
<td>11</td>
<td>1.64</td>
<td>0.67</td>
</tr>
<tr>
<td>Implementer</td>
<td>53</td>
<td>21</td>
<td>26</td>
<td>1.63</td>
<td>0.68</td>
</tr>
<tr>
<td>External environment</td>
<td>66</td>
<td>17</td>
<td>17</td>
<td>1.51</td>
<td>0.78</td>
</tr>
<tr>
<td>Resources</td>
<td>68</td>
<td>28</td>
<td>4</td>
<td>1.36</td>
<td>0.57</td>
</tr>
<tr>
<td>Facilitation processes</td>
<td>92</td>
<td>6</td>
<td>2</td>
<td>1.11</td>
<td>0.37</td>
</tr>
</tbody>
</table>

To produce a quantitative statement of probability of future implementation success, Bridge-It’s mathematical model recodes the factor scores into likelihood ratios and applies the odds version of Bayes' theorem. Instead of weighting and summing the factor scores as would be done in a multi-attribute evaluation, the factor scores are expressed as likelihood ratios and multiplied. The likelihood ratios associated with a “high,” “medium,” or “low” score for each factor were empirically derived in analyses of the expert panel members’ forecasts of likelihood of implementation success for hypothetical profiles (see Bosworth et al., 1999, especially Table 2 which lists the factors, their attributes, and likelihood ratios). The developmental tests of Bridge-It’s internal validity showed a high level of agreement between its forecasts of likely implementation success and the expert panel members’ subjective assessments of likelihood of success for 100 hypothetical profiles (correlation of 0.92, \( p < 0.001 \)). An initial external evaluation of Bridge-It’s utility in forecasting implementation status for TTPI school programs at mid-course (i.e., 2 years after start-up) showed the system was successful in 75% of cases in discriminating (\( p < 0.05 \)) at baseline which programs would achieve a criterion level of implementing activity and which would not (Gingiss, Roberts-Gray et al., 2006).

2.3. Measures of implementation status

To assess implementation of tobacco prevention and control practices relative to the school guidelines for best practices to prevent tobacco use and addiction, principal and health coordinator surveys were designed to correspond to the 2000 SHEP-TM developed by the Centers for Disease Control and Prevention (CDC). The rationale stated at the national level for the tobacco module survey items is to measure the extent to which schools follow the school guidelines to achieve the Healthy People 2010 Objective 27-11 of creating smoke-free and tobacco-free schools (USDHHS, 2000). The surveys include questions to measure each of the multiple components of the school guidelines.
guidelines: policy, instruction, curriculum, training, family involvement, tobacco cessation efforts, evaluation.

The edition of the survey for the school principal included nine questions about development and enforcement of school policy on tobacco. The questions asked how often the policy on tobacco use by students is enforced in school buildings, on school grounds, in school vehicles, and at school-sponsored events (question 1); whether tobacco advertising to sponsor school events is prohibited and whether students are prohibited from wearing or carrying tobacco-brand name apparel and merchandise (questions 2 and 3); how often actions (e.g., referred to legal authorities) are taken when tobacco use policy is violated by students and faculty or staff (questions 4 and 5); which persons (e.g., school administrators, teachers, bus drivers) are responsible for reinforcing school policy that prohibits smoking by students and faculty and staff (questions 6 and 7); by what means (e.g., written in student handbook) policy prohibiting use of tobacco is communicated to students and to their families (questions 8 and 9). The survey questionnaire also asked how much implementing activity there was at the school for each component of the *school guidelines*. This latter item was included in the pre-implementation and mid-course assessments.

The health coordinator edition of the implementation status survey included questions about tobacco use prevention education (TUPE), program-specific training for teachers, cessation support for students and for faculty and staff, and assessment of prevention programs. Provision of instruction and curriculum is assessed in 11 questions that asked: (1) whether TUPE is provided through classroom teachers and, if so, (2) at what grade levels TUPE is provided; (3) whether teaching of TUPE is required (4) how many teachers provide TUPE in the classroom; (5) number of classroom lessons dedicated to TUPE; (6) which classes include TUPE; (7) whether TUPE is provided through non-classroom programs or activities; (8) which persons outside the classroom (e.g., school nurse) are involved in providing TUPE; (9) from what curricula TUPE lessons were taken; (10) which sources of materials (e.g., a commercially developed student textbook) were used to provide TUPE; and (11) what instructional methods (e.g., group discussions) were used to provide TUPE.

Training was assessed with questions that asked whether faculty had received and would like to receive staff development for TUPE, and whether faculty would like to receive staff development on specific methods for providing TUPE (e.g., teaching students of various cultural backgrounds). Questions about family involvement asked in what ways family (parents/guardians) of students have been involved in tobacco use prevention. Questions about tobacco cessation efforts asked whether information has been provided to students and/or faculty and staff for help quitting tobacco use. The questionnaires were piloted with 16 school administrators, health teachers, and graduate education students to assess time for completion, clarity and completeness of directions, vocabulary level, answer sheet format, clarity of questions and answers, and content completeness. A small number (6%) indicated that some changes were indicated in the vocabulary level and answer sheet format, respectively. Those changes were made. Content validity was checked in consultation with 13 of the 20 Texas Regional Educational Service Center health coordinators. Although the CDC has not published reliability assessments for the SHEPTM, the questions are similar to those used in the CDC School Health Policies and Programs Study (SHPPS) in 2000, for which reliability has been demonstrated (Brener, Kann, & Smith, 2003).

To measure quality of adherence to the *school guidelines*, a multi-attribute scoring procedure was developed. Development of the scoring rubrics occurred in consultation with school health experts at national, state, and regional levels. For each survey item, a score of 3, 2, or 1 was assigned for response options indicative of high, partial, or low level of adherence to the *school guidelines*. Adherence is here defined as the quality of being faithful or showing fidelity to the *guidelines*. A percentage score was calculated based on the maximum possible points compared to actual points assigned. Obtained percentage scores from 67 to 100 were accepted as indicating high level of adherence to the *school guidelines*; partial adherence was indicated by percentage scores from 34 to 66; and low adherence was indicated by percentage score less than 34 (Boerm, Gingiss, & Roberts-Gray, 2002).

To provide a measure of quantity of implementing activity, we extracted the survey items that asked about amount of implementing activity for the various components of the *school guidelines*. The raw score response scales for these items were Likert-type ranging from 1 for “not active at all” to 4 for “extremely active” in implementing the given component.

2.4. Analyses

Descriptive statistics were applied to analyze the pre-implementation and mid-course assessments in order to describe school capacity to implement the *school guidelines*. *Bridge-It*’s mathematical model was applied to forecast likelihood of implementation success. Descriptive statistics also were used to analyze the follow-up data collected nearly 4 years after the start of the TTPI to describe two aspects of achieved implementation status: quality of adherence to the *school guidelines* and quantity of implementing activity. We used repeated measures analysis of variance to evaluate changes in capacity from pre-implementation to midcourse and changes in amount of implementing activity from pre-implementation to follow-up. Analyses were conducted using SPSS for PC (Version 11).

To test the hypothesis that the *Bridge-It* model provides valid predictors of future implementation status, we used the pre-implementation and mid-course measures of
capacity as predictors and the follow-up measures of implementation status as criterion variables in regression analyses.

To test the hypothesis that Bridge-It’s multi-factor approach for measuring capacity has utility because different factors are predictive of different aspects of achieved implementation status, we used analysis of variance to compare the pattern of results obtained in predicting quantity of implementing activity versus predicting the quality of adherence to the school guidelines.

3. Results

3.1. Capacity to implement the school guidelines

Forty-six percent of schools were identified as having at least medium likelihood of achieving the quality and quantity of implementation needed for success. The average overall forecast on Bridge-It’s 100-point scale representing the Bayesian probability statement about likelihood of achieving implementation success was low (mean = 23.07, standard deviation 28.15). Expressed as the average across the factor scores (1 = low, 2 = medium, 3 = high), the overall score was 1.64 with a standard deviation of 0.45.

Examination of capacity data for the schools that participated in both the pre-implementation and the mid-course assessment showed a positive increase from pre-implementation to mid-course in the score for external environment (t = 3.12, p < 0.01), but no other significant changes in capacity. There was no significant change in the overall capacity score. We looked in particular for changes in scores for resources and for implementer characteristics because the initial capacity building activities provided through the TTPI included training and guidance for implementers and also provided materials and reimbursed schools up to $2000 for expenses associated with implementing their programs. Inspection of the data sets indicated the factor scores were arrayed in the desired direction but not significantly improved. Average score on the aggregate score scale (1 = low, 2 = medium, 3 = high) for implementer characteristics at pre-implementation was 1.61 ± 0.83 compared with 1.96 ± 0.88 at mid-course; and 1.36 ± 0.56 and 1.46 ± 0.58 at pre-implementation and mid-course respectively for the resources factor. At pre-implementation and again at mid-course, the average score for implementer characteristics was toward the middle of the scale (i.e., “medium”) while the score for resources fell on the bottom third of the scale (i.e., “low”).

Compatibility of the TTPI programs with needs, culture, and history of the school was the only capacity factor for which the average score was in the medium to high range (2.15 ± 0.88 on the scale where 1 = low, 2 = medium, 3 = high). The majority of schools obtained medium or high scores for compatibility (68%), external leadership (57%), innovation characteristics (57%), and school-based leadership (54%). Less than half of the schools obtained medium or high scores for implementer characteristics (47%), external environment (34%), and resources (32%). The factor with the least positive profile was facilitation processes. Only four of the 47 schools (8%) obtained a medium or high score for this factor. These data are displayed in Table 2.

Inspection of the data to identify items for which the survey responses most often were below the scoring criterion for counting the attribute as a positive contribution to the factor score (i.e., response codes less than 4 or 5) indicated the following specific constraints on capacity to implement: lack of time in the resources factor, TTPI being low among the principal’s priorities in the school-based leadership factor, lack of specific implementation skills in the implementer characteristics factor, lack of support from parents and community members in the external environment, and all of the items in the facilitation processes factor. Constraints indicated in responses to the facilitation processes items included a lack of formal implementation plan, lack of access to multiple sessions of implementer training, lack of on-going coaching and technical assistance, lack of monitoring and feedback, and lack of two-way channels of communication.

3.2. Implementation status—quality of adherence to the school guidelines

At the follow-up almost 4 years after the start of the TTPI, 30 of the 47 schools (64%) had implementation status categorized as high quality adherence to the school guidelines for best practices to prevent tobacco use and addiction. None of the schools obtained overall scores that fell on the lower third of the 100-point adherence scale. The average overall score across the school guidelines was 70.11 ± 13.08.

The components for which high scores most often were recorded were development or establishment of school policy regarding tobacco use (94% or 44 of the 47 schools), instruction for students on tobacco use prevention (74%), enforcement of tobacco use policy (60%), and teacher training for tobacco prevention (53%). Family involvement was the only school guidelines component with substantial numbers of schools obtaining a low score—43% or 20 of the 47 schools obtained a low score for adherence to guidelines regarding family involvement in tobacco use prevention. These data are displayed in Table 3.

3.3. Implementation status—quantity of implementing activity

Level of implementing activity across the components of the school guidelines was low to moderate. Average score across components was 2.49 with standard deviation 0.58 (1 = no activity at all to 4 = extremely active).

Data displayed in Table 4 show that enforcing school policy on tobacco use was the school guidelines component where schools most often were extremely active (68% of
Table 3
Percent of TTPI schools (n = 47) shown by level of quality of adherence to the various components of the school guidelines assessed nearly 4 years after start of the TTPI AND average adherence scores across schools

<table>
<thead>
<tr>
<th>Component</th>
<th>Low (%)</th>
<th>Medium (%)</th>
<th>High (%)</th>
<th>Mean</th>
<th>Std. dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Establish policy</td>
<td>0</td>
<td>4</td>
<td>94</td>
<td>84.99</td>
<td>10.61</td>
</tr>
<tr>
<td>Instruction and curriculum</td>
<td>13</td>
<td>9</td>
<td>74</td>
<td>78.72</td>
<td>25.94</td>
</tr>
<tr>
<td>Enforce policy</td>
<td>11</td>
<td>20</td>
<td>60</td>
<td>78.72</td>
<td>20.76</td>
</tr>
<tr>
<td>Teacher training</td>
<td>13</td>
<td>34</td>
<td>53</td>
<td>77.22</td>
<td>25.67</td>
</tr>
<tr>
<td>Cessation support</td>
<td>14</td>
<td>40</td>
<td>46</td>
<td>67.50</td>
<td>22.64</td>
</tr>
<tr>
<td>Program assessment</td>
<td>19</td>
<td>62</td>
<td>19</td>
<td>55.28</td>
<td>16.38</td>
</tr>
<tr>
<td>Family involvement</td>
<td>43</td>
<td>43</td>
<td>14</td>
<td>53.61</td>
<td>22.99</td>
</tr>
</tbody>
</table>

Table 4
Percent of schools (n = 47) shown by quantity of implementing activity for the components of the school guidelines at follow-up nearly 4 years after the start of the TTPI AND average score across schools

<table>
<thead>
<tr>
<th>Component</th>
<th>Not active (%)</th>
<th>Low activity (%)</th>
<th>Moderate activity (%)</th>
<th>Extremely active (%)</th>
<th>Mean score</th>
<th>Std dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enforce policy</td>
<td>6</td>
<td>9</td>
<td>17</td>
<td>68</td>
<td>3.48</td>
<td>0.89</td>
</tr>
<tr>
<td>Instruction for students</td>
<td>4</td>
<td>13</td>
<td>62</td>
<td>21</td>
<td>3.00</td>
<td>0.72</td>
</tr>
<tr>
<td>Program assessment</td>
<td>21</td>
<td>30</td>
<td>38</td>
<td>11</td>
<td>2.38</td>
<td>0.95</td>
</tr>
<tr>
<td>Teacher training</td>
<td>23</td>
<td>40</td>
<td>28</td>
<td>9</td>
<td>2.22</td>
<td>0.91</td>
</tr>
<tr>
<td>Cessation support</td>
<td>25</td>
<td>43</td>
<td>21</td>
<td>11</td>
<td>2.19</td>
<td>0.99</td>
</tr>
<tr>
<td>Establish policy</td>
<td>38</td>
<td>30</td>
<td>19</td>
<td>13</td>
<td>2.07</td>
<td>1.05</td>
</tr>
<tr>
<td>Family involvement</td>
<td>28</td>
<td>45</td>
<td>4</td>
<td>28</td>
<td>2.02</td>
<td>0.76</td>
</tr>
</tbody>
</table>

The majority of schools were moderately or extremely active providing instruction to students (TUPE). Cessation support, establish policy, and family involvement were the components with the least amount of implementing activity. Only 32% of the schools were moderately or extremely active in implementing these aspects of school programs to prevent tobacco use and addiction.

The moderate level of implementing activity at follow-up represents a systematic but modest increase from the amount of activity reported during the pre-implementation phase ($t_{34} = 7.24, p < 0.05$, partial $\eta^2 = 0.18$). The total score at pre-implementation was $17.09 \pm 4.41$ on a scale with lower bound of 8 and upper bound of 32 compared with $19.71 \pm 4.69$ at the follow-up. The school guidelines components with statistically significant increase in implementing activity were instruction for students on tobacco use prevention (TUPE) ($t_{34} = 3.43, p < 0.01$), family involvement in tobacco policy and programs for students ($t_{34} = 3.43, p < 0.01$), support for student cessation ($t_{34} = 2.28, p < 0.05$), and assessment of prevention programs ($t = 2.10, p < 0.05$). There was no change in amount of implementing activity for the policy component. The majority of schools were reported to be “extremely active” in enforcing school tobacco policy at both pre-implementation and follow-up.

3.4. Capacity scores as predictors of implementation status

To test the hypothesis that the Bridge-It system provides valid predictors of future implementation status, regression analyses were conducted using the Bridge-It scores as predictors and implementation status measured nearly 4 years after start-up of the TTPI as criterion. The overall Bridge-It score generated with the Bayesian model for combining data across all eight of the capacity factors was a significant predictor of both quality of adherence to the school guidelines ($\beta = 0.39, p < 0.01$) and quantity of implementing activity ($\beta = 0.33, p < 0.05$). The overall Bridge-It score also was a significant predictor of quantity of implementing activity for the tobacco prevention instruction component ($\beta = 0.39, p < 0.01$) and quality of adherence to components regarding teacher training for TUPE ($\beta = 0.30, p < 0.05$), cessation support ($\beta = 0.40, p < 0.01$), and assessment of tobacco prevention programs ($\beta = 0.37, p < 0.05$).

To test the hypothesis that Bridge-It’s multi-factor approach for measuring school capacity for program implementation has utility because different factors are predictive of different aspects of implementation status, regression analyses were conducted with stepwise entry of the scores for each of the eight capacity factors as predictors. The school-based leadership factor was a significant independent predictor of quality of adherence to the school guidelines ($\beta = 0.45, p < 0.01$), whereas the facilitation processes factor predicted quantity of implementing activity ($\beta = 0.42, p < 0.05$).

The analyses were repeated using only the data for the 28 schools that participated in all three of the evaluation studies. The overall Bridge-It score generated with the Bayesian model for combining data across all eight of the capacity factors at baseline in 2000 tended toward significance in predicting both quality of adherence to the school guidelines ($\beta = 0.35, p = 0.070$) and quantity of
implementing activity ($\beta = 0.37$, $p = 0.055$). The school-based leadership factor was a significant independent predictor of quality of implementation ($\beta = 0.43$, $p < 0.05$). The resources factor predicted quantity of implementing activity ($\beta = 0.52$, $p < 0.01$).

In summary, schools that had higher capacity to implement obtained significantly higher scores for quality of adherence to the school guidelines and for quantity of implementing activity at follow-up nearly 4 years after the start of the TTPI than did schools that had lesser capacity scores. To illustrate these findings we categorized the overall capacity scores as “high,” “medium,” or “low” and used capacity as a fixed factor in analyses of variance with the measures of implementation status as dependent variables. We categorized overall capacity as “low” when the score was on the lower third of the scale (score 1–1.60 averaged across the factors), “medium” for the middle third of the scale (scores 1.61–2.39), and “high” for the upper third of the scale (scores 2.40–3). The data displayed in Table 5 shows capacity at pre-implementation and/or mid-course was a significant effect in explaining the scores obtained at follow-up for quality of adherence to the school guidelines and for quantity of implementing activity.

### 4. Discussion

Secondary analyses of evaluations conducted for school programs in the TTPI supported hypotheses about the validity and utility of the Bridge-It system for assessing school capacity to implement new health and education programs. In a prospective design, the Bridge-It survey tool and Bayesian model produced forecasts that reliably predicted scores obtained at the 4-year follow-up of overall amount of implementing activity and quality of adherence to the Guidelines for School Health Programs to Prevent Tobacco Use and Addiction (school guidelines, CDC 1994, 2004). These positive results are confirming of the initial field test of Bridge-It’s predictive validity that showed the pre-implementation forecasts accurately predicted which schools would achieve criterion level of implementing activity 2 years after start-up of the TTPI (Gingiss, Roberts-Gray et al., 2006).

The current study also supported the hypothesis that Bridge-It’s multi-factor approach for measuring capacity has utility because different factors were predictive of quality versus quantity of implementation. The school-based leadership factor was useful in predicting quality of adherence to the school guidelines whereas the facilitation processes factor and the resources factor were useful as independent predictors of amount of implementing activity.

These findings are consistent with conclusions reached by others that different measures of implementation status provide different perspectives on implementation (Dane & Schneider, 1998; Dusenbury et al., 2003). The measures of implementation status used in the current study, in fact, provided somewhat different statements about success in achieving desired levels of implementation. Nearly 4 years after start-up of the TTPI, approximately two-thirds of the schools met a priori criteria for high level of quality of adherence to the school guidelines. On average, the overall score for quality of adherence was moderate to high. The average score for amount of implementing activity, however, was low to moderate. Only about half of the schools were reported to have at least moderate implementing activity across the multiple components of the school guidelines.

Finding facilitation processes, resources, and school-based leadership factors were that independently had utility in predicting implementation status confirms findings reported in the original development of the Bridge-It system. Bridge-It’s mathematical model assigns substantially greater weight to facilitation processes, resources, school-based leadership, and implementers than to the other four of the eight factors (Bosworth et al., 1999). Finding in the current study that the Implementers factor made no significant independent contribution in predicting implementation status supports conclusions drawn by other researchers that although implementer training is essential, training alone is not sufficient to ensure high fidelity implementation (Dusenbury et al., 2003; Nachmias, Mioduser, Cohen, Tubin, & Forkosh-Baruch, 2004; NIRN, 2005).

Bridge-It is proving to be a useful tool for helping to understand school capacity and offers options for planning and evaluating capacity building initiatives. An as yet untested component of the Bridge-It system is item-by-item feedback about options for leveraging identified strengths and taking action to mitigate detected constraints. In the current study, for example, compatibility of the TTPI...
programs with needs, culture, and history of the school was identified as a common strength and being low among the principal’s priorities was a frequent constraint. The options suggested in Bridge-It’s feedback in response to low or medium score for priority assigned by the principal are: “Strive to place the program high on the list of priorities at the school by providing credible evidence of ways in which it is compatible with identified needs, culture, and structure of the school. Provide information about ease or simplicity of program implementation, and supply evaluation data or other evidence of its beneficial effects. Seek and publicize program endorsements by school and district administrators, parents, and respected local or national opinion leaders.” Bridge-It’s item-by-item suggestions are consistent with and expand guidelines recommended by others for helping schools to increase intensity and quality of program implementation (e.g., Payne et al., 2006).

We expect use of Bridge-It in the early phases of introducing new programs into schools can help program developers and administrators to tailor resource allocation, training, and other support to build schools’ capacity for implementing recommended best practices and evidence-based health and education programs. A pre-implementation assessment showing, for example, the majority of schools that have “adopted” an evidence-based program have no formal plan for implementing and/or no procedures for monitoring its progress could help the program developer or cross-site administrator to make important decisions about adding to or adjusting content of the program “package” to support successful replication of the program in a non-research setting.

5. Limitations

Conclusions drawn from the current study are limited by its relatively small sample size and uneven participation by the schools in the series of evaluation studies from which the datasets were drawn. Although we found no significant differences in characteristics (e.g., school size and type), pre-implementation capacity, or follow-up implementation status between schools included and those that did not meet eligibility criteria for the current analyses, the small sample size relative to the large amount of variability in the scores limited our ability to evaluate relative contributions of the capacity factors as predictors of implementation status. To get a sample size marginally large enough for multiple regression analyses, we included schools for which the capacity assessment was mid-course rather than pre-implementation. That decision limited our ability to easily map the logic of the predictive model, but it allowed us to begin testing hypotheses about validity and utility of the Bridge-It system.

Another limitation was our use of only survey-based measures of program implementation. Observations and interviews are recommended as strategies for measuring qualities of implementation of school programs (Dusenbury, Brannigan, Hansen, Walsh, & Falco, 2005; Lillehoj, Griffin, & Splot, 2004; Resnicow et al., 1998). Although limited, the measures used here have the advantage of being simple to administer, easy to score, and meeting statistical demands for testing Bridge-It’s survey-based Bayesian model.

The current study presents data only for programs to prevent tobacco use and data only for schools in East Texas. In practice, the Bridge-It system has demonstrated its utility for providing sponsors of school health centers a basis for prioritizing the delivery of scarce training and technical assistance resources (Gingiss & Engel, 1995) and eliciting suggestions for improving a program for elementary schools to protect youth from problems with use of alcohol (Bell, Bliss, & Padget, 2005). We expect Bridge-It will perform equally well for informing the dissemination and implementation of a wide range of school-centered health and education programs across the broad spectrum of types and locations of schools because of its firm foundation in research and theory (see Bosworth et al., 1999; Gingiss, Roberts-Gray et al., 2006).

6. Lessons learned

Implementation is a bridge between a school program and its impact on students and their families (e.g., Kalafat, Illback, & Sanders, 2007). But implementation success often is elusive. The current study, like many others (e.g., Hahn et al., 2002; Elliott & Mihalic, 2004; Payne et al., 2006), showed only about half of schools with high levels of quality and quantity of implementation. The lesson reinforced here is that ample time and money for local capacity building should be included in any statewide or national school-centered initiative (Adelman & Taylor, 2003; Elias, Zins, Graczyk, & Weissberg, 2003; Elliott & Mihalic, 2004; Fullan, 2005).

Capacity analysis can play an important role in helping to bridge the gap between expected and achieved benefits of introducing evidence-based practices into schools. The Bridge-It survey tool and Bayesian model for capacity analysis provided forecasts that reliably predicted quality of adherence and quantity of activity to implement best practices for school programs to prevent tobacco use and addiction. The school-based leadership, facilitation processes, and resources factors and the overall measure that combined data about these factors with data about implementers, external leadership, external environment, innovation characteristics, and compatibility of the program with the school were reliable predictors of implementation status measured nearly 4 years after startup of school programs in the TTPI. The lesson learned is that multi-attribute assessment of the situation is a more useful approach than a narrow focus on expected benefits of the program and/or capabilities and commitment of the implementers.

Integration of Bridge-It, or comparable multi-attribute tools, into the planning and evaluation of school-centered programs can increase understanding of factors that
influence implementation and provide guidance for capacity building. Additional research and practice applications still are needed, however, to validate the system with alternative health and education foci, increased sample size, and alternative measures of implementation status. Field trials and additional testing also are needed to ascertain the extent to which the provision of support responsive to Bridge-It’s forecasts has the desired effect of increasing the quality and quantity of implementation and sustainability of multi-component school-centered health and education programs.

Acknowledgment

This study was conducted as part of research sponsored by the Texas Department of State Health Services (TDSHS) under contract 74660013992.

References


Texas Department of State Health Services, TDSHS (2001). *Texas Tobacco prevention initiative: Infrastructure and baseline data*. Austin, TX: Texas Department of Health.