ORIGINAL PAPER

Effects of the Strategic Prevention Framework State Incentives Grant (SPF SIG) on State Prevention Infrastructure in 26 States

**Robert G. Orwin · Alan Stein-Seroussi ·
Jessica M. Edwards · Ann L. Landy ·
Robert L. Flewelling**

Abstract The Strategic Prevention Framework State Incentive Grant (SPF SIG) program is a national public health initiative sponsored by the U.S. Substance Abuse and Mental Health Services Administration's Center for Substance Abuse Prevention to prevent substance abuse and its consequences. State grantees used a data-driven planning model to allocate resources to 450 communities, which in turn launched over 2,200 intervention strategies to target prevention priorities in their respective populations. An additional goal was to build prevention capacity and infrastructure at the state and community levels. This paper addresses whether the state infrastructure goal was achieved, and what contextual and implementation factors were associated with success. The findings are consistent with claims that, overall, the SPF SIG program met its goal of increasing prevention capacity and infrastructure across multiple infrastructure domains, though the mediating effects of implementation were evident only in the evaluation/monitoring domain. The results also show that an initiative like the SPF SIG, which could easily have been compartmentalized within the states, has the potential to permeate more broadly throughout state prevention systems.

R. G. Orwin (✉) · A. L. Landy
Westat, Rockville, MD, USA
e-mail: robertorwin@westat.com

A. Stein-Seroussi · J. M. Edwards · R. L. Flewelling
Pacific Institute for Research and Evaluation, Chapel Hill,
NC, USA

Introduction

Background

The Strategic Prevention Framework (SPF) is a synthesis of a variety of empirically driven planning and implementation models, such as Getting to Outcomes (Center for Substance Abuse Prevention (CSAP)–National Center for the Advancement of Prevention (NCAP), 2000; Imm, Chinman, Wandersman, Rosenbloom, Guckenburg, & LeisSanta, 2007), that emerged in the prevention field over the decade preceding the SPF SIG launch. It also drew on lessons learned from the original State Incentive Grant (SIG) program, in which 44 states and territories were funded between 1997 and 2005. Related developments included a focus on risk and protective factors as a unifying descriptive and predictive framework (CSAP, 2002) and the accumulation of evidence identifying specific prevention interventions that were (or were not) effective in preventing or reducing substance use, the purpose of which was to increase the number of funders requiring or at least encouraging grantees to select evidence-based interventions.

The SPF model consists of five steps: (1) assess population needs, resources, and readiness to address needs and gaps, (2) mobilize and/or build capacity to

address needs, (3) develop a comprehensive strategic plan based on assessment results, (4) implement evidence-based prevention programs, policies, and practices, and (5) monitor, evaluate, sustain, and improve or replace those that fail (CSAP, 2004). The underlying hypothesis of the initiative is that faithful implementation of these SPF steps, with continuing attention to cultural competence and sustainability, will build state and community substance abuse prevention capacity and lead to the selection and implementation of effective and appropriate prevention strategies. This will, in turn, result in reduced consumption levels and substance-related problems in the population.

Another notable characteristic of the SPF model— not present in the original SIG program—is its emphasis on using epidemiological and other data to help identify prevention priorities and allocate prevention resources from a population-based perspective to ultimately achieve population change. It particularly informs that part of planning devoted to deciding which priorities related to consumption and consequence to target, and how to allocate resources to communities to best meet those targeting goals. The SPF SIG represents the first U.S. federal grant initiative in substance abuse prevention that *required* states to engage in data-driven strategic planning and to have their written plans approved by the federal government *prior to* releasing program funds to communities to address substance abuse and related priorities at the local level (Orwin, Edwards, Buchanan, Flewelling, & Landy, 2012). The data-driven planning requirement came with two important tools to support implementation, which if properly utilized were also expected to boost infrastructure. The first was the inclusion in the grant of support for the creation and maintenance of a state epidemiology workgroup (SEW) with the requisite data acquisition, analysis and interpretation skills as well as a diverse membership to assure cultural competence. The second was the provision of intensive, high-quality technical assistance (TA) by a CSAP TA contractor or other partner. This TA included guidance documents and the creation of a publically accessible data base containing a variety of state and county consumption and consequence indicators for every state in the country (Flewelling, Birckmayer, & Boothroyd, 2009), as well as workshops and training materials to further articulate issues and strategies relevant to the organization, analysis, and

interpretation of the epidemiological data for assessment and monitoring purposes.

Assessing Infrastructure Development in the SPF SIG Cross-Site Evaluation

Twenty-four states and two territories received 5-year SPF SIG grants in the fall of 2004 (Cohort I) and summer of 2005 (Cohort II).[1,2] A team of investigators from Westat, the Pacific Institute for Research and Evaluation (PIRE), and The Mayatech Corporation conducted a national cross-site evaluation of Cohort I and II program implementation and outcomes at the state and community levels. The evaluation addressed all three CSAP goals, though the work reported in this paper is confined to Goal 1 (infrastructure development).[3]

While much of substance abuse prevention occurs at the community level, each state's prevention infrastructure can influence the types of prevention policies and programs that are implemented, the populations that are targeted, the level of resources and expertise that are available to support and guide implementation, and the extent to which such efforts are monitored or evaluated. Components of strong prevention infrastructure that have been identified in the literature include: a unifying theory of prevention and a logic model for implementing the theory, effective leadership, collaboration among organizations and agencies, a set of consistent working definitions of key terms (for example, "evidence-based prevention"), comprehensive and effective strategies that affect individuals and environments, monitoring and evaluation, workforce development (e.g., training or accreditation), cultural competence, marketing, and sustainability (Fulbright-Anderson, Kubisch, & Connell, 1998; Johnson, Hays, Center, & Daley, 2004; National Research Council and Institute of Medicine, 2009; Neal, Altman, & Burritt, 2003; Rosenbloom, Leis, Shah, & Ambrogi, 2006; SAMHSA, 2003). Measurement systems have been

---

[1] For simplicity of presentation, the term "state" is used to refer to both states and territories.

[2] Additional awards (Cohorts III, IV, and V) have been made in subsequent years, but are not part of this cross-site evaluation.

[3] A detailed description of the full evaluation design can be found at https://www.spfsig.net/public_general/ShowDocuments. asp?category=20&Category_type=PublicGeneral.

developed and used to assess core capabilities in public health systems, for example the Centers for Disease Control and Prevention (CDC) National Public Health Performance Standards Program (CDC, 2008). However, no well-established measurement tool existed prior to our study that was specifically designed to quantitatively assess states' substance abuse prevention infrastructure, either for the purpose of describing the foundation of prevention systems, measuring changes in infrastructure over time, or assessing the impact of infrastructure on prevention-related outcomes.

To bridge this gap and enable us to address Goal 1, we developed the State-Level Infrastructure Instrument to assess state prevention systems with respect to their infrastructure domains. The instrument was designed to secure information concerning (a) the dimensions of prevention infrastructure as measured at the level of the state agencies responsible for coordinating and funding substance abuse prevention efforts, and (b) whether and to what extent changes in infrastructure occurred during, and resulted from, the state's implementation of the SPF SIG grant. An earlier paper (Piper, Stein-Seroussi, Flewelling, Orwin, & Buchanan, 2012) addressed issues related to dimensionality. Based on extensive interviews with prevention stakeholders in all 26 Cohort I and II states during the early phase of grant implementation, we found that states and territories varied substantially with respect to their organizational structure and each of six functional domains: strategic planning; data systems; workforce development; use of evidence-based programs, policies, and practices (EBPPPs); evaluation and monitoring; and cultural competence. Across the six domains, states scored highest on data systems and lowest on strategic planning. We observed positive associations among the domains, indicating that states with higher capacity on one domain generally have higher capacity on others as well.

The present paper addresses the change question. As described above, we examine changes in state prevention infrastructure that occurred during implementation of the SPF SIG grants and the extent to which such changes could reasonably be attributed to the grant program as a whole. We also examine the degree to which infrastructure change appeared to be mediated by state-level implementation of the SPF model.

Methods

The main data source for assessing state-level infrastructure change was two rounds of interviews conducted with state agencies using the State-Level Infrastructure Instrument (described below). To support the analyses of infrastructure change, we conducted additional interviews to assess the implementation of the five steps, supplemented by review and coding of state strategic plans. We describe each of these data sources below.

Assessment of State-Level Prevention Infrastructure

As noted previously, no well-established measurement tool existed prior to our study that was specifically designed to quantitatively assess the components of states' substance abuse prevention infrastructure. We developed the state infrastructure instrument to measure these constructs through an interview protocol. Because the goal of the grant was systemic statewide change, we designed the instrument to ensure that the attributes assessed were reflective of the state prevention system as a whole, exclusive of any capacities that were in place solely for the purpose of implementing the SPF SIG project.

Round 1 Instrument Development

We developed and pilot tested the instrument in an extended iterative process (see Piper et al., 2012, for details). This included:

- Pre-development open-ended interviews with stakeholders to: (1) ensure that the identified infrastructure domains had face validity; (2) identify additional prevention domains missing from the first typology; (3) learn from professionals how each domain could be operationalized; and (4) establish domain standards that reflect highly developed prevention infrastructure.
- Development, stakeholder review, revision, piloting, and further revision.
- Utilization of a subcommittee of SPF SIG state evaluators to develop a scoring guide.

It is worth noting that while improving infrastructure capacity was a stated goal of CSAP, no requirements to improve specific infrastructure elements

Table 1 Contents of the SPF SIG State Prevention Infrastructure Interviews by domain

Domain	Topics covered
Organizational structure	Level of cooperation, coordination, collaboration among the bureaucratic entities controlling the major prevention funding sources
	Degree to which levels of the system (state-level agencies, regional entities, local entities and prevention providers) work together
Strategic planning	Existence of a strategic plan and (if a plan exists) the plan development process
	Availability of resources for plan development and implementation and level of support among state decision-makers, state agency leaders, sub-state leaders and the local workforce
	Impact of the strategic plan on the system and the extent to which the plan is used to set prevention priorities, goals, and objectives
	Existence of a state plan for addressing cultural competence[a] (Round 2 only)
Data systems	Types of data available and the level of disaggregation of the available data (county, regional, or state) for generating population estimates
	Existence of a state-level group that manages and coordinates the data collection
	Availability of necessary data management and analysis expertise
	Data collection procedures (Round 1 only)
Workforce development	Requirements for workforce competencies and procedures for assessing adequacy and needs of its prevention workforce
	Existence of a workforce development plan
	Workforce development activities implemented in the state
	Cultural competence of workforce development activities[a] (Round 2 only)
Use of evidence-based programs, policies, and practices (EBPPPs)	Consistency of definitions of EBPPPs across statewide agencies and from top to bottom among communities, counties, school districts and other intermediate entities
	Extent to which the use of EBPPPs is emphasized across the state's system and is supported by resources
	Cultural competence of workforce development activities[a] (Round 2 only)
Cultural competence	Existence of a state plan for addressing cultural competence (Round 1 only)
	Extent to which the state provides resources to support culturally-competent prevention interventions (Round 1 only)
Evaluation and monitoring	Resources and expertise available within the system for evaluating and monitoring prevention efforts
	Evaluation requirements within the system
	Use of evaluation and monitoring data
	Degree to which evaluation and monitoring information are clearly communicated between state and sub-state entities

[a] Cultural competence items from Round 2 are in italics

were specified in the grant announcements. However, certain activities were mandated pursuant to that goal, the most salient being the establishment of the SEW, which would directly impact the data systems domain as well as support improvements in strategic planning and evaluation/monitoring capacity. The general lack of requirements or specificity about elements needing improvement was in large part why the evaluation team began the instrument development process by interviewing stakeholders from grantee states about their own goals for state-level infrastructure development, supplemented by a review of the pertinent literature.

The final version used in the first round of interviews (Round 1) consisted of 113 open- and closed-ended questions in seven domains, as shown in Table 1. Before fielding the instrument, we held a 1-day training of all interview teams to ensure their understanding of the instrument and the coding scheme, and to increase the reliability of the scoring

process. At the end of the training, we assessed agreement among team members based on example responses generated through role playing by mock respondents. Inter-rater agreement was 80 % or higher across the items within each domain.

The timing of the Round 1 infrastructure interviews presented a challenge to the evaluation. The time needed to develop a new instrument and then obtain approval from the U.S. Office of Management and Budget (OMB) meant that the interviews—once begun—would not represent a true "baseline" assessment, especially for those states ready to move quickly through the five SPF steps after the grants were awarded. To address this challenge, we incorporated open-ended retrospective questions at the end of each section of the interview. We asked respondents what their state system looked like prior to the SPF project and, if there were changes, how the SPF SIG had contributed to the current status of the system.

Round 2 Modifications

We modified the protocol for the second round (Round 2) of interviews in a number of ways. First, we altered the approach to cultural competence. In Round 1 there was a separate eight-question cultural competence domain. However, our analysis of the Round 1 data showed that the cultural competence domain was only weakly correlated with the other domains. Given guidance from CSAP that cultural competence should crosscut the five SPF steps, we believed it was appropriate to view cultural competence as systemic, in that states' cultural competence efforts would be reflected across all domains of the prevention system rather than being a separate component of the system. The Round 2 interviews still included seven questions on cultural competence but we modified the structure so that they were embedded in the interviews on the use of EBPPPs, strategic planning and workforce development domains.

Second, we pared the protocol from 113 to 93 questions based on an analysis of the variability and utility of the Round 1 answers. Specifically, we deleted items that did not clearly reflect the spectrum of low to high state-level prevention capacity or did not seem to yield clear responses.

Third, we combined the data systems and the evaluation/monitoring domains into one interview session because in Round 1 the same people invariably attended each session. Consequently, we reduced the number of interview sessions from seven to five.

Data Collection

Prior to the interviews we contacted each grantee project director or coordinator to obtain names and contact information for potential respondents nominated by the project director as "domain experts." We conducted separate telephone interviews for each domain with one to five state experts participating in each interview. The Round 1 interviews took place between December 2006 and August 2007 and Round 2 took place between December 2008 and June 2009. Round 1 generated 182 interviews and Round 2 produced 130 (recall that the number of domains had decreased from seven to five). We used recordings, when necessary, to confirm respondents' comments. We then produced narrative summaries of each interview and sent them to SPF SIG project directors and coordinators for their comments and corrections of factual inaccuracies.

Domain Scoring

Following each interview, the interview teams coded responses using scales created for each domain. Although some interview items were purely descriptive and open-ended, most were designed to be coded in one of several ways. Some items were coded as yes/no, for example, "Does the state prevention system have a written strategic plan for substance abuse prevention?" Others were coded for periodicity, e.g., "How often is the strategic plan revised or updated?" Still others were more qualitative in nature and used to rate levels of infrastructure capacity. For example, one item asked "How would you characterize the level of support for the state strategic plan among various stakeholders?" Respondents were not asked to use a rating scale, but instead were asked semi-structured questions with prompts (e.g., "support among state decision-makers," "support among mid-level program managers," and "support among the prevention workforce"). Answers were used to create a scale score of 1 ("no" or "low" prevention capacity), 2 ("moderate prevention capacity"), or 3 ("high prevention capacity").

Using these methods, the interview teams generated scores reflecting the level of prevention infrastructure
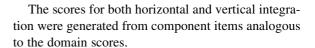
capacity for each of the domains. Most of the individual scores making up a domain score were based on information from multiple questions or components. Domain scores were determined by averaging the component scores to yield a score potentially ranging from one to three.

System Integration Definitions and Scoring

Based on a review of the open-ended interview responses, we also developed indices of horizontal and vertical integration from items spread throughout the interview protocol.

Horizontal Integration refers to the level of cooperation and coordination across the state-level agencies and organizations responsible for substance abuse prevention. The underlying assumption is that a higher degree of horizontal integration leads to more effective prevention efforts across the state. Examples of horizontal integration include (a) agreements among prevention agencies on fundamental operating standards, such as having a statewide plan covering multiple agencies; (b) standard requirements for staff competencies; and (c) consistent policies and definitions. We compiled horizontal integration scores from 13 interview items related to cohesion within the statewide systems (five items from evaluation/monitoring, five from strategic planning, and one each from data systems, workforce development, and EBPPPs).

Vertical Integration refers to how well the various levels of the prevention system (state, regional, and local entities involved in a state's prevention efforts) were coordinated. We compiled vertical integration scores from 18 items across the five domains, plus cultural competence (four items from evaluation/monitoring, four from strategic planning, two from data systems, one from workforce development, three from EBPPPs, and four from cultural competence). In general, the questions asked how state and local prevention agencies related to one another regarding: (a) decision-making within the system; (b) the collection, distribution and interpretation of data for planning; (c) agreement on what it means for a program, policy, or practice to be evidence-based; (d) the availability of state support for evaluation; (e) support for selecting EBPPPs; and (f) coordination among state agencies of their requirements for local prevention programming.

The scores for both horizontal and vertical integration were generated from component items analogous to the domain scores.

Baseline Infrastructure Scores

State-level "baseline status" is the first source of influence in the SPF SIG logic model of impact. We therefore created a measure of baseline infrastructure using the open-ended retrospective questions that were included at the end of each domain of the Round 1 infrastructure interviews. These questions simply asked whether the activities that had been discussed in that domain had occurred prior to the SPF SIG. The intent was to get some indication of the infrastructure that existed prior to the initiative, since it was not possible to conduct the infrastructure interviews before the grants began. Responses to these questions were synthesized across domains to construct a single ratio score that represents an overall pre-SPF infrastructure score. The SPF SIG states' baseline infrastructure scores ranged from .06 to .94, with a mean of .60 and a standard deviation of .24. This indicates that considerable variation existed among the states in prevention infrastructure prior to the SPF SIG, validating the need to control for this variation in subsequent outcome analyses.

Assessment of State-Level SPF Implementation

In addition to asking whether SPF funding led to improvements in state prevention infrastructure, we also asked what accounted for variation in performance across the SPF SIG states. We expected that, all else being equal, higher levels of state-level implementation of the SPF framework should yield greater gains in prevention infrastructure. To explore this question we examined correlations between implementation and infrastructure scores and also employed simple regression analyses to control for differences across states in infrastructure levels prior to SPF SIG. These analyses made use of measures that summarized state-level SPF implementation, as well as the baseline infrastructure scores to control for pre-SPF differences.

Implementation Instrument Development and Data Collection

Separate from the infrastructure interviews, all 26 SPF SIG states participated in additional interviews

assessing the implementation of the SPF. We conducted these interviews by telephone with state SPF SIG project directors, evaluators, SEW chairs and, in some states, additional staff identified as knowledgeable about the interview topics. The development of the SPF implementation interview instrument and the protocol for the first round of interviews took place between January 2005 and February 2006. We developed the instrument on the basis of site visits to SPF SIG states, reviews of documents including the SPF SIG Request for Applications and published literature, expert reviews, and pilot tests with six states. We conducted the first round of interviews between March and August 2007. Preparation for the second round of interviews took place during the first half of 2008. This included making some adjustments to the instrument to improve clarity and avoid redundancy, as well as adjusting the time periods to which some questions referred. The second round was conducted between October 2008 and June 2009.

The content of the SPF implementation interview instrument addressed steps 1, 2, 4, and 5 in the SPF as well as cultural competence, sustainability, and state-level contextual factors that might affect substance abuse, its consequences, and the implementation of SPF projects. The interview included a set of standard questions asked of all states/territories and also questions that were tailored for each grantee based on the state strategic plan. Step 3 was assessed separately, as described below.

Interview Coding and Scoring

Prior to the first round of interviews, we developed a coding form that included both categorical and ordinal response categories. Both the interviewer and the note taker recorded the interview responses on the interview forms and then, shortly after the conclusion of the interview, independently coded the items on the coding form. They then met to discuss their coding decisions, reach consensus regarding any differences of opinion, and complete a single coding document that contained their consensus codes.

During and after the second round of interviews, we developed plans for producing implementation scores for the SPF steps. This included a scoring document that built on the coding form by specifying criteria, or anchors, for assigning scores on a 3-point scale (1 = low to 3 = high implementation) for the

interview items. Scoring of the interviews emphasized information from the second round of interviews in combination with some selected information from the first round. The second round was emphasized because it provided the most recent and comprehensive information about states' implementation activities and future plans.

Strategic Plan Coding

The creation and CSAP approval of state strategic plans completed the SPF's third step. The submission, review, and eventual approval of a state's plan were critical milestones in the SPF SIG implementation sequence, because CSAP approval of the plan was a prerequisite for the release of community funding, which represents 85 % of the grant funds. In the strategic plan, the state justified its selection of target priorities and community funding plan, based on the data-driven model.

Rather than interviewing state SPF SIG project directors and evaluators about their SPF strategic plans, the cross-site evaluation team reviewed the approved plans directly and assessed the strength of the linkage between needs assessment results (Step 1) and decisions about substance use-related priorities (Step 3) and funding allocation decisions, as described in approved plans. We developed the strategic plan coding protocol after thoroughly reviewing CSAP's *Guidance for Developing the State Strategic Plan* (2005) and eight approved plans. The protocol was modified as necessary throughout the coding process to include new codes and topic areas.

Team members reviewed and coded all 26 approved strategic plans over the course of 7 months during 2007. To assess the strength of the linkage states made between needs assessment results (Step 1) and decisions about substance use-related priorities (Step 3), we used the following scale anchors to score each state's priorities:

3 High—Process is transparent and decision is fully supported by relevant evidence.

2 Medium—Process is transparent but decision is only partially supported by relevant evidence *or* Decision appears fully supported by evidence but the process is not fully transparent.

1 Low—Process by which data were used to support priority selection is not transparent *or* Process is transparent but decision is not supported by relevant evidence.

0 No state or community-level priorities were identified.

A similar procedure was employed to rate funding allocation decisions. For further details on methods for coding state strategic plans, see Orwin et al. (2012).

Summary Implementation Scores

We created summary scores for each of the five SPF steps as well as cultural competence and sustainability. With the exception of Step 3, these scores were computed by combining information from multiple interview items. The Step 3 score was based on coded information from the strategic plans. Finally, we created two overall summary implementation scales. The *basic* summary scale was computed by taking the mean of the five SPF step scores and an *augmented* summary scale was computed by taking the mean of the five SPF steps and the scores for cultural competence and sustainability.

Results

Changes in Domain Scores, Integration Scores, and Organizational Structure from Round 1 to Round 2

Domain Scores

As detailed in the "Methods" section, we averaged scores across domain items to create overall domain scores, ranging from 1.0 (no or minimal capacity) to 3.0 (high capacity) for each of the five domains. Table 2 displays the Round 1 and Round 2 scores. As shown, there were statistically significant gains on three domains (strategic planning, workforce development, and EBPPPs) and nominal gains on a fourth (data systems). Given the low power associated with *N* = 26, differences are reported as statistically significant if *p* < .10. For both rounds of data collection, states achieved the highest mean scores on the data systems and EBPPP domains, and the lowest mean scores for strategic planning (with evaluation/monitoring tying strategic planning for lowest mean score in Round 2).

A closer look at the components of these scores provides some insight into the specific areas in which states improved their capacity. In strategic planning,

Table 2 Differences in Prevention Infrastructure Domain Scores between the two rounds of SPF SIG state infrastructure interviews (*N* = 26)

Domain	R1 Mean (SD)	R2 Mean (SD)	*t*(25)
Strategic planning	1.79 (.57)	1.95 (.57)	1.83*
Data systems	2.05 (.47)	2.14 (.42)	1.54
Workforce development	1.82 (.42)	2.02 (.42)	3.21**
EBPPPs	2.04 (.49)	2.24 (.47)	3.50**
Evaluation and monitoring	1.97 (.51)	1.95 (.51)	−.20

* *p* < .10, paired *t* test, two-tailed

** *p* < .05, paired *t* test, two-tailed

more states had strategic plans for substance abuse prevention at Round 2 (19 of 26 states, ten of which served as plans for multiple state agencies) than at Round 1 (14, only three of which served as plans for multiple state agencies), and there seemed to be enhanced mechanisms for linking state planning with sub-state planning. In workforce development, states were more likely to report at Round 2 having increased their formal structures related to workforce development, including enhanced workforce competency requirements (including certification) and a mechanism for assessing their workforce needs. States also reported a greater number of workforce development opportunities and greater access to those opportunities at Round 2. For EBPPPs, states at Round 2 were more likely to report having consistent criteria for defining EBPPPs and requiring their service providers to implement EBPPPs.

We also examined the extent to which individual states moved to higher or lower levels of capacity (or stayed the same) for each domain. Results are displayed in Table 3. For all domains except evaluation/monitoring, at least 40 % of the states that were characterized by low or moderate capacity at Round 1 increased their capacity level at Round 2. In contrast, no more than 25 % of states that were at high or moderate capacity at Round 1 decreased their capacity at Round 2. Disaggregating this finding by domains yields the following:

- Of the twelve states that were low capacity in strategic planning in Round 1, six increased to moderate capacity at Round 2.

Table 3 Changes in SPF SIG State Prevention Infrastructure Domain Capacity Scores ($N = 26$)

Domain	Capacity at Round 1	Capacity at Round 2				Percent of low or mod states (R1) that increased (R2)	Percent of high or mod states (R1) that decreased (R2)
		Low	Mod	High	Total		
Strategic planning	Low	6	6	0	12	50 %	NA
	Moderate	2	3	3	8	38 %	25 %
	High	0	1	5	6	NA	17 %
	Total	8	10	8	26	45 %	21 %
Data systems	Low	0	5	0	5	100 %	NA
	Moderate	2	9	3	14	21 %	14 %
	High	0	0	7	7	NA	0 %
	Total	2	14	10	26	42 %	10 %
Workforce development	Low	3	5	1	9	67 %	NA
	Moderate	3	7	3	13	23 %	23 %
	High	0	1	3	4	NA	25 %
	Total	6	13	7	26	41 %	24 %
EBPPPs	Low	3	1	1	5	40 %	NA
	Moderate	0	9	6	15	40 %	0 %
	High	0	0	6	6	NA	0 %
	Total	3	10	13	26	40 %	0 %
Evaluation and monitoring	Low	3	5	0	8	63 %	NA
	Moderate	4	8	0	12	0 %	33 %
	High	0	0	6	6	NA	0 %
	Total	7	13	6	26	25 %	22 %

Low capacity = scores of 1.0–1.66; moderate capacity = scores of 1.67–2.33; high capacity = scores of 2.34–3.0

- All five states that were low capacity in data systems at Round 1 increased to moderate capacity at Round 2.
- Of the nine states that were low capacity in workforce development at Round 1, six increased to at least moderate capacity at Round 2, of which one increased to high capacity.

Even in the evaluation/monitoring domain (which did not improve significantly overall), five of the eight states that were at low capacity at Round 1 improved to moderate capacity at Round 2. Thus, there was consistent improvement across all domains among low capacity states.

States that were at moderate capacity at Round 1—and could therefore move in either direction—were more likely to increase to high capacity at Round 2 than decrease to low capacity in the EBPPP domain (of the 15 moderates at Round 1, six increased and none decreased), but only slightly more likely to increase to high capacity at Round 2 than decrease to low capacity in most domains. The latter was true for strategic planning (of the eight moderates at Round 1, three increased and two decreased) and data systems (of the 14 moderates at Round 1, three increased and two decreased). The opposite was true for evaluation/ monitoring, where 4 of the 12 moderate states decreased at Round 2 and none increased. For workforce development, an equal number of the 13 moderate capacity states at Round 1 increased and decreased (three each).

Integration Scores

As described in the "Methods" section, we developed indices of horizontal and vertical integration from items spread throughout the interview protocol. Horizontal integration refers to the cooperation and coordination across the state-level agencies and organizations responsible for substance abuse prevention, whereas vertical integration refers to how well the various levels of the prevention system—state, regional, and local entities—work together. As seen

in Table 4, both types of systems integration increased from Round 1 to Round 2, and the increases were statistically significant. A look at component items suggest that the increase in horizontal integration seems to have been driven by increases in state-level groups that meet to integrate prevention across agencies, as well as cross-agency strategic plans and criteria for evidence-based prevention. The increase in vertical integration seems to have been driven by increases in state support to sub-state agencies regarding the selection and implementation of EB-PPPs, sharing of epidemiological data for sub-state planning, and guidelines/requirements for sub-state entities such as workforce competency requirements and criteria for defining EPPPPs.

We also examined patterns of movement to higher or lower levels of horizontal and vertical integration in individual states (see Table 5). For horizontal integration, 6 of the 24 states that scored low or moderate at Round 1 increased their level at Round 2, whereas 1 of the 12 states that scored high or moderate at Round 1 decreased their level at Round 2. For vertical integration, 10 of 21 states that scored low or moderate at Round 1 increased their level at Round 2, whereas of

2 of 22 that scored high or moderate at Round 1 decreased their level at Round 2. States scoring moderate on vertical integration at Round 1 were more likely to increase than decrease at Round 2, with 7 of the 17 moderate states increasing to high and two decreasing to low. The effect was less pronounced for horizontal integration, where one of the ten moderate states increased and none decreased.

Organizational Structure

Several structural components of the states' prevention systems were quite consistent across the states. For example, all 26 SPF SIG states have Single State Agencies (SSAs) that are responsible for coordinating substance abuse prevention services. The SPF SIG project was housed in the SSA in 24 of the states, and 22 states had entities encouraging and facilitating integration of substance abuse prevention efforts among state level agencies (e.g., Departments of Education and Juvenile Justice).

Other components, in contrast, varied from state to state, and changed over the course of the grant. First, eight of these states at Round 1 had "super-agencies," or entities with the authority to make decisions about statewide programs, policies, or resource allocation. Eleven had coordinating bodies: their intent was to help integrate substance abuse prevention efforts, but they had no authority to set policies, create programs, or allocate resources. The remaining seven states had neither a super-agency nor a coordinating body for substance abuse prevention. The number of states with super-agencies did not change between Rounds 1 and

Table 4 Differences in Prevention Infrastructure Integration scores between the two rounds of SPF SIG state infrastructure interviews ($N = 26$)

Integration type	R1 Mean (SD)	R2 Mean (SD)	t(25)
Horizontal integration	1.69 (.57)	1.85 (.57)	2.85**
Vertical integration	2.07 (.47)	2.23 (.42)	2.54**

** $p < .05$, paired t test, two-tailed

Table 5 Changes in SPF SIG States' Prevention System Integration Scores ($N = 26$)

Domain	Level at Round 1	Level at Round 2				Percent of low or mod states (R1) that increased (R2)	Percent of high or mod states (R1) that decreased (R2)
		Low	Mod	High	Total		
Horizontal integration	Low	9	5	0	14	36 %	NA
	Moderate	0	9	1	10	10 %	0 %
	High	0	1	1	2	NA	50 %
	Total	9	15	2	26	25 %	8 %
Vertical integration	Low	1	3	0	4	75 %	NA
	Moderate	2	8	7	17	41 %	12 %
	High	0	0	5	5	NA	0 %
	Total	3	11	12	26	48 %	9 %

Low = scores of 1.0–1.66; moderate = scores of 1.67–2.33; high = scores of 2.34–3.0

2, but the number of states with a coordinating body increased from 11 to 14.

Second, the number of SPF SIG states with regional entities increased from 14 to 15. The regional bodies were an integral part of the state's substance abuse prevention system, providing TA and training to community organizations and local prevention providers.

Finally, the number of states with a line item (i.e., general revenue funds) in their state budgets for substance abuse prevention increased from 14 to 16. The creation of a budget line item is arguably one of the most significant structural changes of those described, as it reflects a state's commitment to substance abuse prevention beyond its sole reliance on federal funding.

We also examined whether the presence of these structures (a superagency, a regional structure, or a budget line item for substance abuse prevention) was associated with domain and integration scores. We found that the budget line item was the only structure associated with these scores. As shown in Table 6, states with a line item in their budgets for substance abuse prevention efforts were more likely to have higher scores on strategic planning, data systems, workforce development, horizontal integration, and vertical integration.

State-Level Implementation as a Mediator of Infrastructure Change

This section addresses the degree to which infrastructure change appeared to be mediated by state-level implementation of the SPF model. This was examined with simple regression models that control for differences across states in infrastructure levels prior to the implementation of the SPF SIG (i.e., at baseline). Because of the small N at the state level, it was not feasible to enter all the individual implementation scores as separate terms in the models, because doing so would have yielded highly unstable coefficients and unreliable significance levels. Therefore, for data reduction purposes, two implementation summary scores were constructed and tested separately. As described in the "Methods" section, the first or *basic* summary scale was constructed by summing the five steps scores and dividing by 5. The second or *augmented* summary scale was constructed by adding

Table 6 Mean differences in Prevention Infrastructure Domain and System Integration Scores for SPF SIG states with ($N = 16$) and without ($N = 10$) a prevention line-item

Domain	With line-item Mean (SD)	Without line-item Mean (SD)	$t(24)$
Strategic planning	2.16 (.44)	1.62 (.63)	2.56**
Data systems	2.29 (.40)	1.92 (.37)	2.36**
Workforce development	2.15 (.43)	1.82 (.33)	2.06*
EBPPPs	2.31 (.43)	2.12 (.54)	1.01
Evaluation and monitoring	1.98 (.46)	1.91 (.61)	.35
Horizontal integration	1.97 (.36)	1.64 (.48)	1.84*
Vertical integration	2.37 (.35)	2.00 (.48)	2.22**

* $p < .10$, two-tailed

** $p < .05$, two-tailed

the cultural competence and sustainability scores to the basic summary scale and dividing by 7.

Models

Three different analysis models of state-level prevention infrastructure outcomes were tested each representing different strengths and assumptions.

1. $(R2 - R1)_i = \alpha + \beta_{(BL)}x_{(BL)i} + \beta_{(IMP)}x_{(IMP)i}$
2. $R2_i = \alpha + \beta_{(BL)}x_{(BL)i} + \beta_{(R1)}x_{(R1)i} + \beta_{(IMP)}x_{(IMP)i}$
3. $R2_i = \alpha + \beta_{(BL)}x_{(BL)i} + \beta_{(IMP)}x_{(IMP)i}$

where R2 is the Round 2 infrastructure score; i is an index variable for state ($i = 1$ to 26); α is a constant for the intercept term; R1 is the Round 1 infrastructure score; $\beta_{(BL)}$ is the influence of baseline infrastructure; $\beta_{(IMP)}$ is the effect of implementation; and $\beta_{(R1)}$ is the effect of R1.

Model 1 examines the effect of implementation $\beta_{(IMP)}$ on infrastructure change, controlling for the influence of baseline infrastructure $\beta_{(BL)}$. It assumes that impact is reasonably captured by the R2 − R1 change score. It also assumes that the differences among states in their capacity to change (because some start out higher on the scale and therefore have less room to move upward) are handled by their baseline status. A potential vulnerability with a change score outcome is reduced reliability (because the change score takes on the measurement errors in both single scores).

Model 2 investigates the effect of implementation on the Round 2 infrastructure score, controlling for the influence of baseline infrastructure and the Round 1 infrastructure score $\beta_{(R1)}$. This analysis eliminates the vulnerability of the change score based outcome, but potentially introduces a new problem of over-justification. Specifically, the Round 1 infrastructure score would be the ideal baseline covariate for assessing implementation effects if it were a true baseline, but as noted elsewhere, the OMB clearance process delayed the fielding of the Round 1 interviews until after implementation began. Consequently, a partial effect of implementation could be "hidden" in the Round 1 responses, thereby "stealing" from (or attenuating) the observed implementation effect.

Model 3 also examines the effect of implementation on the Round 2 infrastructure score, this time controlling for the influence of baseline infrastructure but not the Round 1 infrastructure score. This eliminates any over-justification, but assumes that relevant baseline differences are fully captured by the retrospective responses that constitute the baseline infrastructure score. To the extent this is not true, the vulnerability to over-justification is replaced by a vulnerability to under-justification.

Data do not exist to test any of these assumptions. For this reason, all three models were run to see if they yield congruent findings, reasoning that if they did, our lack of knowledge of the "truth" about the different assumptions would be less important because the results were robust to violations.

Findings and Interpretation

We ran each model for each of the five domains plus vertical and horizontal integration, crossed by two versions of implementation (the basic and augmented summary scales). This yielded 42 regression runs in all (3 models × 7 domains × 2 implementation measures). While this number of statistical tests increases the risk of capitalizing on chance, the intent here was not to achieve a particular significance level or confirm a particular hypothesis, but rather to see what, if any, interpretable patterns would emerge that potentially shed light on the role that implementation levels played in advancing infrastructure development in SPF SIG states. Our analyses employed SAS GLM under ordinary least squares (OLS) assumptions.

Model results showed that only the evaluation/monitoring domain was reliably predicted by a state-level implementation summary score. Specific findings follow.

- An increase in the basic implementation summary scale predicted an increase in the evaluation/monitoring outcome when the R2 − R1 difference score was regressed on the predictors (Model 1), and also in both models in which the Round 2 domain score was regressed (Models 2 and 3). In Model 1, a 1-point increase in the summary scale predicted a 1.31 increase in the difference score $[t(23) = 1.79, p < .10]$. In Model 2, a 1-point increase in the summary scale predicted a 1.55 increase in the R2 domain score $[t(22) = 2.35, p < .05]$. In Model 3, a 1-point increase in the summary scale predicted a 1.88 increase in the Round 2 domain score $[t(23) = 2.34, p < .05]$.

- An increase in the augmented summary scale predicted an increase in the evaluation/monitoring outcome under both models in which the Round 2 evaluation/monitoring domain score was regressed on the predictors (Models 2 and 3), but not the model in which the R2 − R1 difference score was regressed (Model 1). In Model 2, a 1 point increase in the summary scale predicted a 1.61 increase in the Round 2 domain score $[t(22) = 2.21, p < .05]$. In Model 3, a 1-point increase in the summary scale predicted a 2.03 increase in the Round 2 domain score $[t(23) = 2.32, p < .05]$.

- Results across model assumptions were generally congruent, in that (1) the regression coefficients for implementation were consistently significant and positive when predicting evaluation/monitoring, over and above effects of baseline infrastructure, and (where applicable) Round 1 infrastructure; and (2) with few exceptions, the regression coefficients for implementation were not significant and showed no particular pattern when predicting the other four domains and the two types of integration.

Discussion

In this paper, we addressed the question of whether the SPF SIG contributed to increases in states' substance abuse prevention systems. To answer this question, we

first needed to ascertain whether the states' prevention infrastructures changed during the life of the SPF SIG. The data gathered through hundreds of structured interviews with state prevention stakeholders suggest that state prevention infrastructure did, indeed, change for the better during the life of the SPF SIG project for Cohort 1 and 2 states. We found the following changes between Round 1 and Round 2 interviews:

First, there were statistically significant increases in three of five infrastructure domains (strategic planning, workforce development, and EBPPP) and nominal increases in a fourth (data systems). There was consistent improvement across all five domains among the lowest capacity states, of which 40–100 % increased their capacity level by Round 2, depending on domain. By contrast, states that were at moderate capacity at Round 1, and could therefore move in either direction, were, with one exception (the EBPPP domain), only slightly more likely to increase to high capacity at Round 2 than decrease to low capacity in Round 2. These data indicate that those states that needed such improvement the most were the most likely to improve their state prevention infrastructure.

Second, with respect to horizontal and vertical integration, both types showed increases. The increase in horizontal integration seems to have been driven by increases in state-level groups that meet to integrate prevention across agencies, as well as cross-agency strategic plans and criteria for evidence-based prevention. The increase in vertical integration seems to have been driven by increases in state support to sub-state agencies regarding the selection and implementation of EBPPPs, sharing of epidemiological data for sub-state planning, and guidelines/requirements for sub-state entities such as workforce competency requirements and criteria for defining EPPPPs. In addition, there was considerable movement from moderate to high capacity on vertical integration but almost no movement from moderate to high capacity on horizontal integration, suggesting that once a certain horizontal integration threshold is reached, more than an SPF SIG grant is needed for further upward movement.

Third, across domain and integration scores, 25–45 % of low or moderate capacity states (that is, among states with the potential to increase) did increase their level of capacity from Round 1 to 2, compared to 0–24 % of moderate or high states that decreased.

Fourth, there were several changes in the organizational structure of some states during the SPF SIG project. Specifically, the number of SPF SIG states with a coordinating body to help integrate substance abuse prevention efforts across state agencies, regional entities to provide TA and training to community organizations and providers, and a line item in their state budgets for substance abuse prevention all increased between Round 1 and Round 2. It appears, then, that substance abuse prevention may be gaining some ground in terms of the integration of prevention services across agencies, availability of regional TA and training, and dedication of state resources to prevention efforts (rather than relying solely on federal sources of funds), and that the SPF SIG may have aided these gains. Neither the presence of a superagency nor the presence of a regional structure was associated with higher domain or integration scores, but the presence of a budget line item was positively related to both domain and integration scores; specifically, states that had a line item in their state budgets for substance abuse prevention efforts were more likely to have higher scores on strategic planning, data systems, workforce development, horizontal integration, and vertical integration. While the causality of this relationship is unknown, it does suggest that a budget line item for prevention is an important marker for state infrastructure system maturity.

Finally, regression models showed that implementation levels consistently predict evaluation/monitoring, over and above effects of baseline infrastructure and Round 1 infrastructure; and that the regression coefficients for implementation were mostly not significant and showed no particular pattern when predicting the other four domains and the two types of integration.

Did the SPF SIG Program Cause the Observed Infrastructure Changes?

The findings are consistent with claims that overall, the SPF SIG program has met its goal of increasing prevention capacity and infrastructure at the state level. They also lend credence to the notion that an initiative like the SPF SIG, which could easily have been compartmentalized within the states, has the potential to permeate broader areas of substance abuse prevention in state systems. Of course, the usual

limitations of observational studies apply to these inferences (e.g., limited ability to make causal attributions, lack of a true baseline, inability to eliminate the possible effects of selection, concurrent historical changes). These states, and possibly some of the non-SPF SIG states as well, may have experienced positive changes in their prevention infrastructures because of a variety of factors such as a better knowledge base about effective prevention practices or more comprehensive overall TA and training from CSAP. From a study design standpoint, it would have been desirable to collect comparable infrastructure data from key informants on prevention in non-funded states, but this was not administratively feasible. Thus, the pre-post design does not allow us to definitively declare that the SPF SIG caused these positive changes.

There is other support, however, from the open-ended interview questions. Many state respondents explicitly indicated that the SPF SIG contributed positively to aspects of their prevention infrastructures beyond the bounds of the grant itself. In particular, respondents reported that planning, use of data to establish priorities, and community/provider training for their Substance Abuse Prevention and Treatment (SAPT) block grants had, indeed, improved as a result of the SPF SIG. They mentioned that the SPF SIG gave them a planning process to follow, increased their awareness of (and knowledge about) the data they processed, and prompted them to increase the capacity of prevention stakeholders across their states.

Evaluation and Monitoring: A Special Case

While most infrastructure domains and both integration types showed positive significant changes, there were some exceptions. Specifically, we did not find overall increases in data systems or evaluation/monitoring—the two domains that are the most directly data-oriented. However, data systems at Round 1 revealed the highest average domain score (and thus had relatively little room to move in a positive direction), yet as shown in Tables 1, 2 they continued to move upwards even if the mean change was not statistically significant.

The mean for evaluation/monitoring, on the other hand did not move at all (see Table 1) and as shown in Table 2, there was virtually no difference between the percent of states that increased and decreased, which is unique among the domains and integration scores.

This finding begs the question of what is different about evaluation/monitoring. That is, are there particular characteristics that differentiate this domain from the others? There are at least two that emerge from the data.

First, an analysis of correlations (not shown) revealed that evaluation/monitoring had the weakest associations with the other domains. Round 2 scores for the other four domains were all significantly correlated with one another, while evaluation/monitoring was only significantly correlated with data systems. Furthermore, evaluation/monitoring change scores were not significantly correlated with change scores from any other domain or integration score—a characteristic unique to evaluation/monitoring. These results suggest that this domain is something of a "lone wolf" among the infrastructure domains, whose pattern of variation and change is determined by a different set of factors. This interpretation is consistent with the fact that evaluation services are often contracted out by state agencies and therefore may be less aligned and coordinated with other functions that are managed and operated internally.

Second, one of those different factors was clearly the implementation strength of the SPF SIG grant itself. As noted above, the regression modeling results showed that only the evaluation/monitoring domain was reliably predicted by variation in state-level implementation; no other domain or integration score was significantly influenced by implementation variance after controlling for baseline infrastructure. This doubly unique pattern of results for evaluation/monitoring—no overall change yet positive change for strong implementers—suggests that the other domains may have been more robust to variation in implementation strength. That is, once a certain implementation threshold was achieved, infrastructure outcomes followed, and additional implementation strength did not raise outcomes further. In contrast, evaluation/monitoring was sensitive to variation in implementation strength—with high implementation leading to positive change and less-than-high implementation leading to no change or even negative change. Although other explanations of what we found are possible, this one is consistent with the data and not uncommon in the prevention literature (Botvin, Baker, Dusenbury, Botvin, & Diaz, 1995; Sloboda et al., 2009). It also suggests that had implementation not been measured and tested, its role in mediating evaluation/monitoring

would have remained hidden, and one might have concluded (erroneously) that the SPF SIG grant failed to affect this important domain.

State-Level Implementation as a Predictor of State-Level Infrastructure Change

Findings related to evaluation/monitoring notwithstanding, it was still somewhat surprising that the implementation scores did not independently predict infrastructure gains in other domains, raising the question of whether one or more methodological artifacts may have been responsible for our findings. Four possible artifacts are examined below: low power, low variability, measurement issues, and over-reducing the independent variable.

- Power is clearly an issue with an N of only 26. Had all of the coefficients approached significance, it would be reasonable to speculate that with more power those effects might have been significant. However, that was clearly not the case, since the other domains were almost perfectly split between positive and negative non-significant effects. That is, across models the implementation coefficients were consistently positive for the strategic planning and data systems domains and for vertical integration, and consistently negative for the workforce development and EBPPP domains and for horizontal integration. Thus the non-significant coefficients appeared to be randomly distributed across these domains, and there is no evidence that a greater N would have yielded greater implementation effects.

- Further, low variability was clearly a factor in the contribution of Step 4 to the summary scales, as our measures of Step 4 had no variation at all. It was also a factor for Step 1, which had very little. Had there been more variation in the implementation scores—which would have meant more states with lower scores—then implementation may have been a stronger predictor. From a programmatic standpoint, the question is somewhat academic, and particularly so if the "threshold" hypothesis raised above is true, insofar as it is good news that implementation of the five steps was high in most states and that most states gained ground on prevention infrastructure.

- Our measures of implementation may have failed to capture everything relevant to predicting infrastructure. However, considerable development time, review, and revision went into the interview protocol (including review by the grantees who would later serve as respondents), augmented by formal interviewer training and extensive state-specific preparation prior to each interview. In addition, the lead interviewer and response coders for the cross-site were retained across both rounds. This not only reduced the likelihood of any interviewer or coder artifacts, but also facilitated a deep understanding of the states and their implementation issues. Consequently, we do not believe that measurement issues significantly biased or attenuated the relationships we observed.

- As noted above, it was not feasible to enter discrete implementation scores for each step as separate terms in the regression models given the small N, so we used summary scores instead. It is possible that with sufficient N to support the simultaneous testing of separate step scores, more evidence of implementation effects would have appeared. It is equally possible, however, that the coefficients would have been reduced to nonsignificance by the inclusion of the baseline infrastructure scores.

In sum, the general lack of a predictive relationship between implementation scores and infrastructure gains (evaluation/monitoring excepted) does not appear to be a methodological artifact, though it likely was constrained by the high level of implementation (and thus low variability) across participating states. We cannot know what the relationship would have been had more states implemented poorly but, and as noted above, the question is academic because that did not occur. Future studies should explore whether state-level implementation predicts the quality of SPF SIG implementation at the community level and, ultimately, the degree of community-level change in substance use and its consequences. These outcomes may be more sensitive to small variations in state implementation.

Limitations

As noted above, delays in obtaining OMB clearance prevented the Round 1 infrastructure interviews from

beginning before some state-level implementation had already occurred. Consequently, responses did not represent true baseline infrastructure status. To obtain more accurate data on baseline status, open ended items at the end of each section of the Round 1 infrastructure interviews asked whether each type of activity discussed in the section had occurred in the state prior to the SPF SIG initiative. These were compiled to form the baseline infrastructure scores, subsequently used in the regression analyses. As a measure of pre-SPF infrastructure levels, these data have the advantage of being based on standardized questions asked of informed respondents nominated by the project director in each state. Nonetheless, we note that responses were retrospective and respondents varied in their exposure to the pre-SPF SIG history of the states. There was also likely to have been considerable variation across states in the way the respondents interpreted the questions, remembered the relevant history, and adopted criteria for deciding whether an infrastructure activity had been in place prior to the SPF SIG.

Additional limitations of both rounds of the infra-structure interviews themselves are also worth noting:

- Because the evaluation was conducted in conjunction with the SPF SIG project, it was sometimes a challenge to maintain respondents' focus on the overall substance abuse prevention system and not just the SPF SIG project. This was particularly the case with small states, whose SPF SIG projects were larger than their SAPT block grant funds and, therefore, became the focus of their prevention systems.
- We note the potential for interview participants to provide socially desirable responses, and believe that a few respondents gave answers they believed would cast them in the best light for the funding source (CSAP). We took several steps to mitigate the influence of social desirability. First, we tried to involve respondents from multiple agencies whenever appropriate, including those with no material interest in the success of the grant (e.g., other state departments that were not the grantee, universities and private research organizations, private, non-profit training centers). Many of these, as it turned out, were quite skeptical of the SPF model and its potential to improve prevention. Second, we attempted to minimize bias though

questions that asked for very specific and objective information, as opposed to respondent opinions and summary judgments.
- The respondents varied in terms of their position in the hierarchy, length of experience, and candor, and this may have affected the consistency of responses. They also varied in number. In some cases up to five people were in the group responding to the questions. That the number in the group affected the responses cannot be ruled out.
- In some cases, the respondents for Round 1 differed from those for Round 2. Thus, when responses did differ, we attempted to determine if the differences were genuine changes in the prevention system or an artifact of our respondents' different perspectives about the system. We attempted via the interview probes and through an extensive post-interview review of the data to try to ensure that changes in the Round 1 and Round 2 scores reflected actual systems changes.

The power to statistically detect pre-post change was limited by the small N (26). Using the standard alpha level of $p < .05$, estimated power to detect a "medium" effect size of .5 (Cohen, 1988) was less than .30. Therefore, we used the more liberal alpha level of $p < .10$. We recognize that we cannot know if the effects detected at $p < .10$ would have been sustained had the N been larger. As it happens, however, several effects were large enough to be significant at $p < .05$ (change in workforce development and EBPPP domains, and both horizontal and vertical integration).

Conclusions and Next Steps

As noted in the introduction, CSAP's goals for the SPF SIG initiative included both building prevention capacity through improved infrastructure and reducing substance abuse and its consequences. The present paper addressed the first of those goals, and the findings suggest that it was largely met, at least at the state level. Across infrastructure domain and integration scores, 25–45 % of low or moderate capacity states (that is, of states with the potential to increase) enhanced their level of capacity during the grant period, which may be compared to 0–24 % of moderate or high states that decreased. There was

consistent improvement across all domains among low capacity states, and across all but one (evaluation/monitoring) among states in general. Changes in organizational structure suggest that substance abuse prevention may be gaining some ground in terms of integration of prevention services across agencies, availability of regional TA and training, and dedication of state resources to prevention efforts (rather than relying solely on federal sources of funds). They further suggest that an initiative like the SPF SIG, which could easily have been compartmentalized within the states, has the potential to permeate broader areas of substance abuse prevention within state systems. This belief was reinforced by the many state respondents who explicitly indicated that the SPF SIG contributed positively to aspects of their prevention infrastructures beyond the bounds of the grant itself, up to and including applying SPF principles to their SAPT block grants.

Data reported elsewhere suggest considerable capacity-building activity at the community level as well.[4] For example, all 26 state administrators required their community partners to conduct and submit a formal needs assessment, develop and submit a data-driven strategic plan, and revise the plan as needed for approval before intervention dollars were released. All these activities were supported by substantial training and TA. Communities were also active in forming new partnerships and strengthening existing ones (e.g., with law enforcement agencies, schools or school districts, media organizations, and youth groups) in the service of building capacity to comprehensively address their priorities. Community level data suggest that, on average, fidelity to the SPF steps has been fairly robust, though considerable variability exists both within and across states. At the intervention level, communities self-reported that the SPF's emphasis on data-driven planning and population-based outcomes had had a major impact on community prevention programming—that is, how programs were selected as well as how they were adapted as needed for implementation. This has been particularly evident in the increased use of environmental strategies. Almost half (44 %) the

interventions implemented by June 2009 were environmental strategies, suggesting that states and communities are taking seriously the SPF SIG goal and challenge of effecting population-level change.

States and communities that embraced the SPF model may also have a smoother transition to the Patient Protection and Affordable Care Act (ACA). The ACA has renewed the country's focus on prevention, emphasizing the need for planning, infrastructure development, and data-based solutions through the articulation of a National Prevention Strategy. The SPF is consistent with the ACA as it offers one tool that policymakers may consider as a means to promote evidence-based prevention decision-making for a variety of health behaviors.

That said, not all states saw improvement, and in three of the infrastructure domains, 20 % decreased in capacity from Round 1 to 2 (see Table 3). Given all the factors that might have contributed, and an N of only 26, it is impossible to do more than speculate as to the reasons. This would be an interesting issue to explore in future studies of prevention infrastructure. Other contextual factors in state prevention systems also merit additional research. For example, the relative importance of leadership and adaptability may be important in facilitating capacity growth, but were not directly assessed in this study. More broadly, further research is warranted on the factors that help to promote and facilitate infrastructure development, together with the barriers that serve to impede its development. Many questions regarding population outcomes and sustainability remain to be answered.

Sustainability

As noted earlier, Cohort I and II states and their funded communities were expected to attend to sustainability from the outset, and sustainability was envisioned as cutting across the five steps of the model. R2–R1 comparisons demonstrated that infrastructure changes were achieved, and other interview responses (in both the infrastructure and implementation interviews) suggest that states took steps to facilitate sustainability. In the implementation interviews, for example, nearly all states reported their intent to maintain elements of the SPF beyond the life of their grants, especially by using the SPF model for statewide prevention planning. This is consistent with responses to the infrastructure interviews regarding application

---

[4] For additional information on community capacity building, see the SPF SIG National Cross-Site Evaluation Phase I Final Report at: https://www.spfsig.net/public_general/ShowDocuments.asp?category=26&Category_type=PublicGeneral.

of SPF principles to the SAPT block grant. In addition, almost all states reported ongoing activities to encourage communities to plan for SPF sustainability. At the same time, they were realistic in their expectations that SPF components would likely be implemented in different and less costly ways after federal funding ended. They also are acutely aware of budgetary pressures from the state and the adverse impact they may have.

References

Botvin, G. J., Baker, E., Dusenbury, L., Botvin, E. M., & Diaz, T. (1995). Long-term follow-up results of a randomized drug abuse prevention trial in a white middle-class population. *Journal of the American Medical Association (JAMA), 273*, 1106–1112.

Center for Substance Abuse Prevention. (2002). *Pathways to effective programs and positive outcomes*. Downloaded from http://eaptus.samhsa.gov/southwest/resources/documents/Pathways-8-03.pdf.

Center for Substance Abuse Prevention. (2004). *Strategic Prevention Framework State Incentive Grants (Short Title: SPF SIG) SP 04-002*. (Initial Announcement).

Center for Substance Abuse Prevention. (2005). *Guidance for developing the state strategic plan*. Washington, DC: Center for Substance Abuse Prevention (CSAP) Substance Abuse and Mental Health Services Administration (SAMHSA).

Center for Substance Abuse Prevention–National Center for the Advancement of Prevention. (2000). *Getting to outcomes*: *Methods and tools for planning, self-evaluation and accountability*. Washington, DC: Center for Substance Abuse Prevention (CSAP), Substance Abuse and Mental health Services Administration (SAMHSA) and National Center of the Advancement of Prevention (NCAP).

Centers for Disease Control. (2008). *National Public Health Performance Standards Program*. Accessed on July 19, 2010 at: http://www.cdc.gov/od/ocphp/nphpsp/essentialphservices.htm.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillside, NJ: Lawrence Earlbaum.

Flewelling, R. L., Birckmayer, J., & Boothroyd, R. (2009). Developing a data resource to support state substance abuse prevention planning: Approaches used in implementing CSAP's strategic prevention framework. *Contemporary Drug Problems, 36*, 387–407.

Fulbright-Anderson, K., Kubisch, A., & Connell, J. (Eds.). (1998). *New approaches to evaluating community initiatives: Theory, measurement and analysis* (Vol. 2). Queenstown, MD: The Aspen Institute.

Imm, P., Chinman, M., Wandersman, A., Rosenbloom, D., Guckenburg, S., & LeisSanta, R. (2007). *Preventing underage drinking: Using getting to outcomes with the SAMHSA Strategic Prevention Framework to achieve results*. Santa Monica, CA: RAND.

Johnson, K., Hays, C., Center, H., & Daley, C. (2004). Building capacity and sustainable prevention innovations: A sustainability planning model. *Evaluation and Program Planning, 27*, 135–149.

National Research Council and Institute of Medicine. (2009). *Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities*. Washington, DC: National Research Council and Institute of Medicine.

Neal, J., Altman, K., & Burritt, S. (2003). *South Carolina toolkit for evidenced-based prevention programs and strategies*. http://www.daodas.state.sc.us/documents/toolkit.htm.

Orwin, R. G., Edwards, J. M., Buchanan, R. M., Flewelling, R. L., & Landy, A. L. (2012). Data-driven decision-making in the prevention of substance-related harm: Results from the Strategic Prevention Framework State Incentive Grant Program. *Contemporary Drug Problems, 39*(Spring), 73–106.

Piper, D., Stein-Seroussi, A., Flewelling, R. B., Orwin, R. G., & Buchanan, R. (2012). Assessing state substance abuse prevention infrastructure through the lens of CSAP's Strategic Prevention Framework. *Evaluation and Program Planning, 35*, 66–77.

Rosenbloom, D., Leis, R., Shah, R., & Ambrogi, R. (2006). *Blueprint for the states: Policies to improve the ways states organize and deliver alcohol and drug prevention and treatment*. Boston: Join Together.

Sloboda, Z., Stephens, P., Pyakuryal, A., Teasdale, B., Stephens, R. C., Hawthorne, R. D., et al. (2009). Implementation fidelity: The experience of the adolescent substance abuse prevention study. *Health Education Research, 24*(3), 394–406.

Substance Abuse and Mental Health Services Administration. (2003). *Summary of "State Substance Abuse Prevention Infrastructure Workgroup" meeting: An internal working document (December 2003)*. Washington, DC: Center for Substance Abuse Prevention (CSAP) Substance Abuse and Mental Health Services Administration (SAMHSA).