Prevention Science (2019) 20:1200–1210
https://doi.org/10.1007/s11121-019-01037-x

Influence of an Implementation Support Intervention on Barriers and Facilitators to Delivery of a Substance Use Prevention Program

Jill S. Cannon[1] · Marylou Gilbert[1] · Patricia Ebener[1] · Patrick S. Malone[2] · Caitlin M. Reardon[3] · Joie Acosta[1] · Matthew Chinman[1]

Published online: 31 August 2019

Abstract

Implementation support interventions have helped organizations implement programs with quality and obtain intended outcomes. For example, a recent randomized controlled trial called Preparing to Run Effective Programs (PREP) showed that an implementation support intervention called Getting To Outcomes (GTO) improved implementation of an evidence-based substance use prevention program (CHOICE) run in community-based settings. However, more information is needed on how these interventions affect organizational barriers and facilitators of implementation. This paper aims to identify differences in implementation facilitators and barriers in sites conducting a substance use prevention program with and without GTO. PREP is a cluster-randomized controlled trial testing GTO, a two-year implementation support intervention, in Boys & Girls Clubs. The trial compares 15 Boys & Girls Club sites implementing CHOICE (control group), a five-session evidence-based alcohol and drug prevention program, with 14 Boys & Girls Club sites implementing CHOICE supported by GTO (intervention group). All sites received CHOICE training. Intervention sites also received GTO manuals, training, and onsite technical assistance to help practitioners complete implementation best practices specified by GTO (i.e., GTO steps). During the first year, technical assistance providers helped the intervention group adopt, plan, and deliver CHOICE, and then evaluate and make quality improvements to CHOICE implementation using feedback reports summarizing their data. Following the second year of CHOICE and GTO implementation, all sites participated in semi-structured interviews to identify barriers and facilitators to CHOICE implementation using the Consolidated Framework for Implementation Research (CFIR). This paper assesses the extent to which these facilitators and barriers differed between intervention and control group. Intervention sites had significantly higher average ratings than control sites for two constructs from the CFIR process domain: *planning* and *reflecting and evaluating*. At the same time, intervention sites had significantly lower ratings on the *culture* and *available resources* constructs. Findings suggest that strong planning, evaluation, and reflection—likely improved with GTO support—can facilitate implementation even in the face of perceptions of a less desirable implementation climate. These findings highlight that implementation support, such as GTO, is likely to help low-resourced community-based organizations improve program delivery through a focus on implementation processes.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11121-019-01037-x) contains supplementary material, which is available to authorized users.

✉ Jill S. Cannon
cannon@rand.org

Marylou Gilbert
marylou@rand.org

Patricia Ebener
pateb@rand.org

Patrick S. Malone
malone@malonequantitative.com

Caitlin M. Reardon
caitlin.reardon@va.gov

Joie Acosta
jacosta@rand.org

Matthew Chinman
Chinman@rand.org

[1] RAND Corporation, 1776 Main Street, Santa Monica, CA 90401, USA

[2] Malone Quantitative, Durham, North Carolina, USA

[3] Center for Clinical Management Research, VA Ann Arbor Healthcare System, Ann Arbor, MI, USA

Trial Registration
This project is registered at ClinicalTrials.gov with number NCT02135991 (URL: https://clinicaltrials.gov/show/NCT02135991). The trial was first registered May 12, 2014.

Keywords Implementation support · Fidelity · Evidence-based program · Community-based, · CFIR · GTO

Background

Research has shown that many evidence-based programs, including substance use prevention interventions delivered in schools and youth-serving organizations (Ennett et al. 2003; Hallfors and Godette 2002; Kumar et al. 2013; Ringwalt et al. 2008; Ringwalt et al. 2009), are often not implemented with fidelity to the program model, resulting in poor outcomes (Durlak and DuPre 2008). It is critical to understand and address these implementation challenges now, as schools and communities struggle to address the opioid epidemic, the dramatic growth in the use of electronic cigarettes among youth (Johnston et al. 2017), the legalization of marijuana, and the rising costs of substance use, which altogether is estimated to cost more than $400 billion annually (U.S. Department of Health and Human Services 2016). Fully implementing an evidence-based program with fidelity (i.e., adherence to and high-quality delivery of the program model components) can be difficult in the best of circumstances, and more so in low-resourced community-based organizations with few staff who are assigned multiple duties and may lack the capacity—defined as knowledge and skills—for evidence-based program implementation with quality (Scaccia et al. 2015). Implementation scientists continue to call for more research in such low-resourced settings (Glasgow et al. 2012).

In the field of prevention of youth problem behaviors, several implementation support interventions have been developed to build the capacity of community-based practitioners to deliver substance use prevention interventions with quality, such as Communities that Care, Promoting School-community-university Partnerships to Enhance Resilience (PROSPER), and Getting To Outcomes (Chinman et al. 2016; Hawkins et al. 2009; Spoth et al. 2007). These interventions take a particular approach, mostly focusing on building the knowledge and skills among community practitioners for the key tasks needed to effectively choose, plan, implement, and self-evaluate evidence-based programs. These implementation support interventions have been successful in enhancing prevention program skills among individual practitioners, improving implementation of evidence-based programs (e.g., adherence to program curricula) and in many cases, also yielding stronger individual youth outcomes (Chinman et al. 2016; Hawkins et al. 2009; Spoth et al. 2007). What is less known however is how these types of implementation support interventions may influence other barriers or facilitators besides individual skills—e.g., leadership support; external pressures such as guidelines and mandates from national headquarters or competition for youth enrollment; or understanding how an evidence-based program's characteristics influence its implementation. This paper describes the influence of an implementation support intervention—Getting To Outcomes (GTO)—on a wide range of implementation barriers and facilitators in low-resourced, community-based settings that are responsible for delivering an evidence-based program to prevent substance use.

This paper uses data from a 29-site, cluster-randomized controlled trial called Preparing to Run Effective Prevention (PREP). PREP evaluated the effects of assignment to the GTO intervention on fidelity of an after-school substance use prevention program called CHOICE being implemented in low-resource community-based settings—Boys & Girls Club. Boys & Girls Clubs are nonprofit organizations that provide youth services ranging from recreation to programs addressing leadership, character education, health and wellness, and academics. While preventing substance use among youth members is compatible with the broader Boys & Girls Club goals, such organizations typically have limited success implementing evidence-based programs (Chinman et al. 2018b). The GTO implementation support is designed to strengthen the knowledge and skills needed to conduct certain program implementation best practices (Livet and Wandersman 2005)—goal setting, planning, evaluation, quality improvement, and sustainment. GTO offers three types of support: (1) the GTO manual (Chinman et al. 2004), (2) face-to-face training, and (3) ongoing, onsite, proactive technical assistance (TA) that is designed to guide practitioners through the GTO practices as applied to their own programs.

Previous research has shown that the GTO intervention can improve multiple components of fidelity (i.e., model adherence and delivery quality) through improved planning and evaluation practices (Acosta et al. 2013; Chinman et al. 2016, 2018a). Similarly, results from the PREP study reported elsewhere show that Boys & Girls Clubs assigned to GTO implemented CHOICE with greater adherence to the program curriculum and with greater delivery quality according to the motivational interviewing model that underlies the CHOICE program (Chinman et al. 2018b). However, we do not know which specific barriers and facilitators GTO may have influenced that could explain the improvement in implementation it yielded.

We used the Consolidated Framework for Implementation Research (CFIR) (Damschroder et al. 2009) as the theoretical

framework to assess facilitators and barriers to CHOICE implementation and help explain the mechanisms for the positive fidelity results found for intervention sites. CFIR was developed as a practical guide for systematically identifying and assessing barriers and facilitators to new program implementation. It includes 39 "constructs" that can influence implementation success, organized around five domains: intervention characteristics, inner setting, outer setting, individuals' characteristics (e.g., the knowledge and skills of individual practitioners), and implementation process. These constructs can facilitate or hinder evidence-based program implementation, depending on how they manifest in organizations, and studies using CFIR have sought to identify relationships between different constructs and implementation effectiveness (Damschroder and Lowery 2013; Damschroder et al. 2017).

In theory, organizations that manifest these constructs more positively would be expected to implement programs more effectively; for example, having greater leadership support for an evidence-based program (i.e., an inner setting characteristic) would result in improved fidelity to an evidence-based program model. It was anticipated that because of GTO's focus on goal setting, planning, evaluation, and quality improvement, it would positively influence constructs in CFIR's implementation process domain: (a) planning; (b) goal setting; and (c) reflecting and evaluating (see Table 1 for a complete list of constructs used). Beyond that expectation, the study is an exploration of what other constructs may have been influenced by assignment to the GTO intervention. It is possible that GTO could help practitioners overcome barriers and leverage facilitators identified by CFIR, for example, by helping program implementers use data and engage leadership to make decisions about whether to support and sustain new programs. CFIR domains address many of the same concepts of organizational readiness for innovations, such as those reviewed by Scaccia et al. (2015), and this study can add to the discussion of readiness based on some of the CFIR constructs. In sum, this study capitalizes on a randomized study that collected rich interview data from practitioners to systematically study how an implementation support intervention may affect potential factors (i.e., CFIR constructs) hypothesized to influence evidence-based program implementation.

Methods

Study Design

PREP was a two-year randomized controlled trial comparing 15 Boys & Girls Club sites that received typical training to implement an evidence-based substance use prevention program, called CHOICE (control group), with 14 Boys & Girls Club sites that received the same CHOICE training, plus the GTO implementation support intervention (i.e., GTO manuals, training, and technical assistance; intervention group). GTO was provided over a two-year period, allowing all sites to deliver CHOICE twice (between May 2014 and April 2016). The study team assessed fidelity (i.e., curriculum adherence, quality of CHOICE delivery, dosage) and the alcohol and drug outcomes of participating middle school youth ages 10 to 14. Details on fidelity, alcohol, and drug outcomes are available elsewhere (Chinman et al. 2018b). In brief, while the GTO intervention sites had better implementation fidelity (e.g., adherence, delivery quality), alcohol and drug outcomes were not different between the intervention and control sites, in part because all youth engaged in very little substance use. Commensurate with the focus of this paper, we conducted CFIR-based semi-structured interviews with multiple staff at each of the 29 Boys & Girls Club sites after the two-year intervention period to assess facilitators and barriers to CHOICE implementation. This research was approved by RAND's Institutional Review Board. No adverse events were reported during the study.

Participants and Setting

All greater Southern California Boys & Girls Clubs were invited to participate, and the recruitment ceased when the above sample was reached. A single Boys & Girls Club often had sites in several locations within the community. The site sample size was justified at 80% power by taking into account the estimated correlation between baseline and follow-up assessments of the site fidelity measures (.5 to .6) and the moderate to large effect sizes for implementation and substance use outcomes expected based on previous GTO (Chinman et al. 2016) and CHOICE (D'Amico et al. 2012) studies, respectively. Each site typically had its own facility and a small number of full- and part-time staff ($n = 7$–10). A sub-set of staff (between 1 and 10; mean = 2.2, median = 3) at each site participated in the study based on their role implementing CHOICE.

Assignment to Study Condition and Baseline Equivalence

We randomized at the club level the 29 Boys & Girls Club sites (i.e., 15 Clubs comprising 29 sites). Seven Boys & Girls Clubs, comprising 14 sites, were assigned to intervention, and eight clubs, comprising 15 sites, were assigned to control using a random number generator. Randomization was done at the club level (versus site) to minimize potential contamination due to communication between sites within the same club system. The principal investigator informed each club about their assignment before study activities commenced.

After randomization at baseline, we conducted a web-based survey with 100% of invited Boys & Girls Club staff involved in CHOICE—typically frontline, part-time staff

Table 1 Consolidated Framework for Implementation Research (CFIR) construct definitions and examples of how scoring was operationalized for the study

Construct definition	Implementation facilitators	Implementation barriers
Innovation characteristics domain		
Relative advantage: perception better or worse than existing programs	Perceived and discussed by staff or leadership as being a better option for programming relevant to age group and subject matter	Perceived to be same as or worse than other drug and alcohol youth programming available
Adaptability: perception of ability to modify program to fit site's needs	Flexibility, inclusivity, and creative use of additional tools	Discomfort with and/or lack of ability to modify program
Complexity: how easy or hard the program is to deliver	Program perceived as easy, short, comfortable to deliver	Program perceived as difficult, overwhelming, and unfamiliar
Outer setting domain		
Needs and resources: extent to which participants' needs are known and prioritized	Good grasp of youths' needs and program adjusted to better suit them	Unfamiliarity with youths' needs or how programming can achieve better results
Inner setting domain		
Networks and communications: informal/formal meetings at site	Frequent meetings, shared information to engage staff with programming	No meetings, discussions, or communications among leadership and staff
Culture: consistency of staffing and site programming	Consistent staffing, little turnover, and programming info is passed on	Large turnover, burden on staff results in inconsistent programming
Implementation climate: general level of staff awareness/receptivity to program	Staff on board with and discuss new programs	No staff awareness and lack of acceptance for new programming
Compatibility: program fits within existing programming, mission, and time frame	Good fit with mission, staff qualities, experience, concerns of community, and timing of program	Not a good fit with staff or community and conflicts with other programming
Relative priority: shared perception of program's importance within site	Other activities planned around new program, program importance is highlighted to community	Scheduling issues and/or lack of interest or lack of perceived need
Leadership engagement: leaders' commitment, involvement, and accountability	Strong leadership, staff reported high commitment	No leadership support, leadership decentralized
Available resources: staff, space, and time	Thorough, periodic review of resources in place	Inconsistent and/or unqualified staffing, no space
Access to knowledge and information: training on the mechanics of the program	Training and program information considered important for accurate delivery	Training did not occur and/or not all staff were trained, programming inconsistent
Implementation process domain		
Planning: pre-implementation strategizing and ongoing program-refining activities	Detailed and ongoing planning process with staff and leadership input	No planning or understanding of what planning is
Formally appointed internal implementation leaders (FIL): individuals responsible for delivering program	Leadership/staff teamwork, evidence of skill, experience, and engagement	Leadership and/or staff have no confidence in FIL, inconsistent staffing, poorly qualified FIL
Innovation participants: youth recruitment, engagement, and retention strategies	Strategies are used to increase youth interest, participation, and attendance	Recruitment issues, lack of youth interest, and inconsistent attendance
Goal setting*: program goals are communicated and acted upon by staff	Goal setting was specific, reviewed, and changed over time to meet needs/expectations	No goals established or discussed
Reflecting and evaluating: quantitative and qualitative feedback about progress and quality of program delivery	Reflecting and evaluation process used to control, shift focus, and make changes	No review of goals nor feedback given either during or after implementation

FIL, formally appointed internal implementation leader. Asterisk denotes an inductively-derived construct

working with youth or site supervisors—on demographic variables and attitudes toward evidence-based programs. Staff had similar demographic makeup between the experimental groups: half (51%) were male; most (59%) were over 25 years old; half (50%) had a four-year college degree or more; and 56% were Hispanic or Latino, 22% were non-Hispanic African-American, and 22% were non-Hispanic White, multiracial, or of other races. Over a third (38%) were full-time.

Attitudes toward evidence-based programs were assessed with the Evidence-Based Practice Attitude Scale (Aarons et al. 2010) to judge the degree to which practitioners would adopt an evidence-based program based on factors such as intuitive appeal, whether it is considered mandatory, willingness to try new interventions, and belief that experience is important. There were no significant differences between the two groups on this measure at baseline (see Chinman et al. 2018b for details of this analysis).

Prevention and Implementation Interventions

CHOICE consists of five 30-min sessions designed to prevent substance use in middle school youth through discussions and role plays about the pros and cons of using alcohol and drugs, social norms about alcohol and drug use, unrealistic beliefs about use, addressing risky situations, and developing coping strategies (D'Amico et al. 2005). The intervention is delivered using motivational interviewing techniques (Miller and Rollnick 2012). The training for CHOICE implementers was 12 hours over 3 days. CHOICE is supported by evidence from two randomized trials that found the program was associated with prevention of the initiation of alcohol use (D'Amico et al. 2012; D'Amico and Edelen 2007).

Using existing staff, all sites were asked to implement CHOICE once a year for two years with a different group of youth each year, between May 2014 and April 2016. For sites assigned to the GTO intervention, Master's level TA providers delivered GTO manuals, conducted face-to-face training on the GTO steps, and provided onsite support prior to CHOICE implementation as well as onsite, phone, and email follow-up support during the two rounds of CHOICE delivery (per site in year 1: $M = 11.17$ h of TA, SD = 3.4; year 2: $M = 14.7$ h of TA, SD = 3.9). The GTO manual contains guidance about how to complete multiple implementation best practices important for evidence-based program delivery—i.e., GTO steps. Most GTO steps contain tools that prompt practitioners to make and record decisions about practices known to support high-quality implementation—i.e., goal setting, program planning, evaluation, quality improvement, and program sustainment. Extensive details about GTO implementation are available elsewhere (Chinman et al. 2018b).

Data Sources

Data collection included semi-structured CFIR interviews with Boys & Girls Club site staff to assess contextual barriers and facilitators to CHOICE implementation in each site. All interviews were conducted by phone by one experienced interviewer and lasted approximately 45–60 min. All respondents agreed to having the interview audio recorded and each recording was transcribed verbatim. For each site, at least two staff were invited to be interviewed about their participation in CHOICE implementation. We attempted to interview both the facilitator who delivered the CHOICE sessions and a site or Boys & Girls Club leader who supervised the facilitator. In most cases, they were interviewed separately but in a few sites, they were interviewed together. Typically, these were the staff who had been trained in CHOICE and/or GTO, but in some cases, staff turned over by the time of these interviews. In total, we conducted 51 interviews—28 were from sites assigned to the intervention and 23 were from sites assigned to the control condition.

These interviews occurred after the completion of the year 2 CHOICE delivery, which was the end of the implementation cycle for PREP and GTO intervention. Interviews took place from March through October 2016 with the aim that they occur within 2 to 4 months after the final CHOICE session. Some of the interview questions asked the respondents to refer back to round one of CHOICE implementation in considering the implementation context. Others asked about plans for a third round of CHOICE delivery, which was at the discretion of the sites.

Analytic Approach

In this paper, we employed a mixed-methods approach that merged quantitative results in a statistical analysis (CFIR ratings) with illustrative qualitative quotes to add further context (Creswell et al. 2011).

Coding: CFIR Interviews and Summary Ratings

The existing CFIR codebook template (https://cfirguide.org/) guided development of the codebook, which consisted of CFIR construct definitions and coding inclusion/exclusion criteria (Damschroder and Lowery 2013). Analyzing the interview content involved both deductive and inductive coding, following Damschroder and Lowery (Damschroder and Lowery 2013). Deductive coding derived codes from CFIR constructs and inductive coding derived codes from the interview data (Damschroder and Lowery 2013). As is common practice for studies using the CFIR, the construct definitions and coding inclusion/exclusion criteria were operationalized and refined over the course of the coding process to meet the specific circumstances of this study. In PREP, this adaptation resulted in a codebook that began as a set of 31 constructs across 4 CFIR domains but was revised to a final set of 17 constructs and 4 domains. Constructs were excluded from analysis for two primary reasons: they were not relevant to the organizational context within the PREP study and thus not asked in the interview, or sites yielded too few responses to be meaningfully analyzed. Similar to Damschroder and Lowery's use of CFIR to assess context within program implementation (Damschroder and Lowery 2013), one domain—Individual Characteristics—was not included and excerpts about individual site staff attitudes, behaviors, and roles were coded to more specific constructs, such as the leadership

and facilitator constructs. In consultation with the CFIR developers, the final codebook included one new inductive construct for *goal setting*, a key implementation process not differentiated from the planning construct in the CFIR process domain. We also note that the *culture* construct definition was operationalized to include statements about staffing consistency. The CFIR *culture* construct is broad and can be operationalized in a number of ways to reflect the study context (see CFIR Research Team undated). In this study, based on statements from Club staff, it was operationalized to reflect a general norm among Boys & Girls Clubs to hire part-time staff who often have short tenures at the site, which affects staffing consistency. Thus, it was meant to assess the *perception* of the impact of this consistency, not *actual* turnover rates. Table 1 provides a list of the final set of 17 constructs and their definitions for this study. See Damschroder et al. (2009) for a detailed description of the full set of CFIR constructs.

The coded interview data were used to make the numerical ratings for the individual constructs under the four CFIR domains included. In general, ratings are determined based on two aspects: (1) valence (positive or negative impact on implementation) and (2) strength (the degree to which implementation was facilitated or hindered, with possible choices being 0, 1, or 2) (Damschroder and Lowery 2013). The valence component of a rating is determined by the influence the coded data exhibits on the implementation process; that is, whether the contextual factor in question was broadly facilitating or hindering to CHOICE implementation. The strength of a rating has several determinants, including level of agreement among respondents, strength of language, use of concrete examples, and stated influence on implementation or another construct. Considering both the valence and the strength of each construct, the rating scale ranges from + 2 (most facilitating) to − 2 (most hindering). A zero rating reflected a neutral (e.g., statements lacking sufficient information to make a clear rating) or mixed influence of the code (e.g., statements indicating equal presence and absence of a construct). A missing code was assigned to reflect when a construct was not mentioned in interviews. Table 1 provides examples of positive and negative indicators for the 17 constructs analyzed in this study.

Each interview transcript was independently coded by a team of two analysts. Following completion of coding, data were aggregated by site and construct and each construct was rated (on a scale from − 2 to + 2) by a team of two research analysts who had received training by CFIR developers. Discrepancies in coding and rating were resolved through a consensus discussion approach. The merging of the ratings and qualitative data was facilitated by creating a matrix (Damschroder et al. 2017) listing Boys & Girls Club sites in columns and CFIR constructs in rows. Each site-construct cell included the CFIR rating and a short summary of qualitative data. Only four constructs had ratings present for all 29 sites, and the other constructs had from 15 to 28 site ratings used in analysis. One analyst was blinded to the sites' outcomes for CHOICE adherence and the second analyst was a PREP TA provider who had knowledge of intervention sites' adherence. Dedoose version 8.0.35 was used to facilitate coding (Dedoose Version 8.0.42 2016). For each site, all the CFIR ratings, illustrative quotes, and the rationale for each rating were summarized in a "site memo." Each site memo was then discussed among the wider research team, which yielded some revisions in the ratings.

Comparison by Study Condition

For each CFIR construct, we calculated mean ratings by study condition. We examined whether any constructs distinguished between intervention and control study conditions. Although the CFIR ratings are on a 5-point scale, in numerous cases, the full range of rating options was not endorsed. Given the sparse and skewed ratings, we transformed the CFIR 5-point rating scale into dichotomized ratings for each construct to indicate positive or negative ratings. To compare study conditions on each of the 17 CFIR constructs included, we conducted Fisher's exact tests of independence on the resulting 2 (condition) X 2 (response) tables. Unlike the chi-squared test of independence, as an exact test, Fisher's does not depend on asymptotic properties and is valid for any sample size (Agresti 2002). We applied the Benjamini-Hochberg correction for false discovery rate across the 17 indicators (Benjamini and Hochberg 1995).

Results

Table 2 presents 17 CFIR constructs, their average ratings across sites by study condition, and the number of sites by intervention or control condition that were categorized as reporting positive or non-positive ratings. The far-right column provides the p value of the 2×2, or positive/non-positive proportions, comparison using Fisher's exact test.

Across all sites, a majority of sites had a positive (i.e., facilitating) rating for seven constructs: *relative advantage* (57% of intervention sites had positive ratings, 77% control), *needs and resources* (100% and 87%), *networks and communications* (78% and 67%), *compatibility* (93% and 82%), *leadership engagement* (71% and 60%), *access to knowledge and information* (64% and 57%), and *formally appointed internal implementation leader (FIL)* (92% and 93%). None of the average ratings for these constructs were significantly different between study conditions.

Positive CFIR Ratings by Study Condition

As anticipated, intervention sites had higher ratings than control sites on several constructs within the *implementation*

Table 2 Differences in construct ratings by intervention or control site condition

Construct	GTO sites			Control sites			*p* value
	Avg. rating	Positive[a] *N* (%)	Non-positive[b] *N* (%)	Avg. rating	Positive[a] *N* (%)	Non-positive[b] *N* (%)	
Relative advantage ($n = 26$)	0.857	8 (57)	6 (43)	0.846	10 (77)	3 (23)	0.420
Adaptability ($n = 26$)	0.462	6 (46)	7 (54)	0.462	7 (54)	6 (46)	1.000
Complexity ($n = 21$)	0.300	5 (50)	5 (50)	0.545	7 (64)	4 (36)	0.670
Needs and resources ($n = 29$)	1.071	14 (100)	0 (0)	1.067	13 (87)	2 (13)	0.483
Networks and communications ($n = 15$)	0.778	7 (78)	2 (22)	0.667	4 (67)	2 (33)	1.000
Culture[c] ($n = 21$)	− 0.917	1 (8)	11 (92)	− 0.111	7 (78)	2 (22)	0.002
Implementation climate ($n = 29$)	0.071	6 (43)	8 (57)	0.067	6 (40)	9 (60)	1.000
Compatibility ($n = 25$)	0.929	13 (93)	1 (7)	0.818	9 (82)	2 (18)	0.565
Relative priority ($n = 28$)	− 0.143	3 (21)	11 (79)	0.214	4 (29)	10 (71)	1.000
Leadership engagement ($n = 29$)	0.786	10 (71)	4 (29)	0.733	9 (60)	6 (40)	0.700
Available resources ($n = 29$)	− 0.214	3 (21)	11 (79)	0.533	10 (67)	5 (33)	0.025
Access to knowledge and information ($n = 28$)	0.786	9 (64)	5 (36)	0.286	8 (57)	6 (43)	1.000
Planning ($n = 25$)	0.769	9 (69)	4 (31)	0.083	2 (17)	10 (83)	0.015
FIL ($n = 27$)	1.250	11 (92)	1 (8)	1.133	14 (93)	1 (7)	1.000
Innovation participants ($n = 26$)	0.071	7 (50)	7 (50)	0.500	8 (67)	4 (33)	0.453
Goal setting ($n = 22$)	0.273	3 (27)	8 (73)	0.000	0 (0)	11 (100)	0.214
Reflecting and evaluating ($n = 24$)	0.857	11 (79)	3 (21)	0.400	4 (40)	6 (60)	0.092

p value denotes test of significant difference in positive and non-positive proportions between study conditions. Significance determined using Fisher's exact tests. After false discovery rate adjustment, *culture* is significantly different between conditions. *FIL*, formally appointed internal implementation leader

[a] Positive denotes construct ratings of + 1 or + 2

[b] Non-positive denotes construct ratings of 0, − 1, or − 2

[c] For *culture* construct only, because of very small positive cell sizes, positive denotes ratings of 0, + 1, or + 2 and non-positive denotes ratings of − 1 or − 2

process domain. Given the study's exploratory nature, we report on findings significant at $p < 0.10$. To provide additional contextual meaning around these findings, we use qualitative data illustrating how sites described the constructs in their settings.

For the *planning* construct, data from the intervention sites showed significantly higher ratings (0.769 average) than control sites (0.083 average; $p = 0.015$). This finding suggests that intervention sites employed more effective pre-implementation planning, including strategizing and refining a program implementation plan as needed. Most of the sites with positive ratings interpreted planning as an activity related to program delivery preparation. Examples include adding CHOICE programming to the site's master calendar, planning to make sure the program fits in and is given appropriate attention, and regularly consulting the plan. Additionally, a higher rating applies to sites that have formally embedded planning practices within site processes—e.g., where both the site leadership and the program's facilitator regularly meet during the implementation period and make revisions if needed. Two staff in different roles at the same site explained how they viewed the planning context:

> "So, I think that the planning process would be incorporated in all aspects of the programming." (Site leader)
> "I would say planning was very important. The planning process, even though it's tedious at times, did help us get ourselves organized and keep us accountable for what we want to accomplish with running Project CHOICE." (CHOICE facilitator)

For the *reflecting and evaluating* construct, the differences in average ratings by condition were significant at the 10% level (intervention, 0.857, and control, 0.400; $p = 0.092$). Data from staff employed at the intervention sites were coded as having higher rates of reflection and evaluation activities. Sites that had positive ratings were likely to review program data and reflect upon it to assess whether implementation changes were needed (e.g., to further engage youth). Higher ratings were assigned to sites where staff reported they were aware of the program's impact. A leader at a site with higher program fidelity noted:

> "I reviewed results with [CHOICE facilitators], how they did, how the group did and then overall…,

> What's working? What's not? Is this something that we think we could expand to the entire organization as opposed to just these two sites, these two employees? … I saw the benefits right away, so we did make a commitment ... to incorporate, continue with Project CHOICE but also GTO."

Intervention sites also had higher average ratings for the inductive *goal setting* construct, although not significantly different. Specifically, only GTO sites had any positive ratings (i.e., + 1) for this construct. Those sites with positive ratings included in their implementation activities a consideration of youth outcomes, not merely program attendance, and building consensus among staff. As explained by one of these sites' staff, it is important to "be able to get data from returning youth and be able to use that data to educate and inform our youth."

Negative CFIR Ratings by Study Condition

In contrast, only the *culture* construct, operationalized here as implementation staff turnover, had a negative average rating for both intervention (− 0.917) and control (− 0.111) sites, indicating that negative ratings were assigned to this construct more often across sites than for other constructs. This was more common for intervention sites than control sites—11 of 12 intervention sites' *culture* ratings were negative compared with 2 of 11 control site ratings ($p = 0.002$). These ratings showed that intervention sites perceived greater barriers to implementation because of turnover among staff implementing CHOICE than control sites. For context, it should be noted that *actual* turnover rates—defined as the percent of sites that had changes in CHOICE implementation staff between round 1 and round 2 implementation—were similar between intervention (71%) and control (67%) sites.

Available resources is another construct for which intervention sites had more negative ratings (− 0.214 average) than control sites (0.533 average; $p = 0.025$), indicating that intervention sites on average faced more barriers associated with organizational resources such as space to implement the program. After the false discovery rate correction, only the *culture* construct showed a significant difference between conditions.

Discussion

As part of a larger randomized controlled trial, this paper compared community-based sites conducting an evidence-based substance use prevention program with and without GTO support on the barriers and facilitators to program implementation measured by CFIR constructs. Intervention sites had significantly higher (more favorable to implementation) average ratings compared with control sites for constructs in the CFIR *process* domain: *planning* and *reflecting and evaluating*. At the same time, intervention sites had significantly lower (less favorable to implementation) ratings on the *culture* and *available resources* constructs, reflecting perceptions of poorer site conditions over the two years of implementation such as the negative effects of staff turnover and limited site resources to facilitate implementation. Intervention and control sites did not significantly differ on any other measured CFIR constructs. This pattern of results could be because the GTO support model directly targets planning and evaluation processes but does not have specific training dedicated to the more general areas of implementation readiness such as organizational leadership or climate for change.

We are unable to say precisely why intervention sites scored more negatively on *culture* than control sites even though sites across both groups faced similar high staff turnover between CHOICE implementation rounds. One consideration could be that intervention staff had a heightened sensitivity to the benefit of staffing consistency after participating in GTO and thus were more likely to comment on it. Given that staff were also trained in and using GTO, there may have been a greater perceived loss when someone left in the intervention group. We note that many control sites either failed to mention the issue of staff turnover or did not indicate it affected CHOICE implementation. The differences might also be related to the specific staff who left positions at intervention sites. For instance, one scenario could be that higher-capacity individuals such as college students were assigned to GTO and CHOICE because of the added responsibilities, so the loss of that staff person is perceived more negatively. Further study of staffing implications related to undertaking GTO support in lower-resourced settings would be helpful.

The CFIR results help elucidate the fidelity findings from the larger randomized trial. Despite the results showing the sites assigned to intervention had lower ratings on *culture* and *available resources* than the control sites, intervention sites had higher CHOICE program fidelity than the control sites (Chinman et al. 2018b). These findings buttress the idea that GTO facilitates improved implementation within existing organizational capacities, even when certain capacities to provide programming may be low. It could be that the implementation support of GTO is more important than some aspects of an organization's overall initial readiness for implementation. In other words, lower readiness factors, at least as measured by these two CFIR constructs, could be overcome by the GTO support. Prior studies using a measure specifically created to assess performance of skills targeted by the GTO support (e.g., goal setting, planning, evaluation) have shown that sites with GTO support perform better in those areas (Chinman et al. 2008; Acosta et al. 2013; Chinman et al. 2016). However, this study adds to that knowledge base by suggesting that GTO can improve areas such as planning and evaluating using a different, non-GTO specific measure (i.e., CFIR).

Youth substance use outcomes in this study were unaffected. These results could be because the rates of substance use and endorsement of risk factors for substance use were very low in this sample (Chinman et al. 2018b). In a previous GTO trial, the target population had more difficulty with the target behavior (engaging in sex, lack of knowledge about condoms), and we found improvement in certain outcomes (Chinman et al. 2018a). Our study indicates that GTO can play a role in improving implementation of an evidence-based program such as CHOICE, and we believe GTO could be used with any similar evidence-based program.

The implications of these findings could be that providing concrete implementation support for planning may be more critical than enhancing certain aspects of organizational readiness. A similar conclusion was reached by Helfrich and colleagues in their evaluation of implementation support for evidence-based workplace health promotion practices (Helfrich et al. 2018). Following Weiner's theory of organizational readiness to change (Weiner 2009), Helfrich and colleagues assessed various aspects of readiness at baseline and 15 months later and no aspect of readiness improved over time, even though implementation efforts did improve for the study group receiving support. They conclude that compared with readiness "constructs, such as commitment, efficacy, and motivation….more instrumental constructs, such as the planning and technical support…may be more important variables in ensuring effective implementation (Helfrich et al. 2018 p. 7)." Clearly, more research is needed given these findings are contrary to the common view that readiness is critical to strong implementation (Scaccia et al. 2015). We note, however, the current study included a limited number of readiness factors compared with other readiness approaches.

This work also contributes to implementation science broadly in that it provides further support for a theoretical framework such as CFIR, but in low-resource community-based settings, which tend to be less represented in implementation research (Glasgow et al. 2012). A systematic review of studies using CFIR found that the great majority were conducted in healthcare systems (Kirk et al. 2016). More work to validate CFIR in community-based settings and to show the role that implementation support interventions like GTO can play in promoting effective evidence-based program delivery is needed given these settings face unique challenges (e.g., culture, resources). Furthermore, unlike most studies employing CFIR (Kirk et al. 2016), we used the application of CFIR to guide both data collection and data analysis, which helped ensure we measured all relevant factors. We also linked implementation factors to program fidelity through comparative analysis of intervention sites (which had higher fidelity ratings) to control sites (which had lower fidelity ratings) using a quantitative ranking of CFIR constructs.

Another area for further research would be to use the CFIR ratings over time to determine how various barriers and facilitators interact with implementation support models. It also would be useful to understand more fully how some of the CFIR constructs may act together in the facilitation of implementation success. For instance, this study showed that, in general, leadership engagement and the formally appointed internal implementation leader had positive CFIR ratings, but we cannot determine if or how they might have also interacted with each other and with GTO TA to improve adherence ratings. Understanding the relationships between CFIR constructs would help identify where community-based programs with limited resources should expend effort to improve implementation success.

Some limitations should be noted. First, the staff interviews were conducted after two years of evidence-based program implementation. Staff were asked to provide retrospective information, but some respondents were not involved in year 1 of the program delivery, so that limited their ability to reflect on the full two-year implementation cycle. Second, we had limitations often found with key informant interviews in that we rely on the information provided by the available informants. In some cases, we had an informant who served two roles (i.e., program facilitator and site leader), so we did not have two separate informants for that site. Third, we interpret our findings as indicating that planning processes changed within intervention sites over the study period and that the organizational climate remained static, but we lack sufficient information to prove that interpretation. We cannot be completely certain of the trajectory of CFIR results. More specifically, construct ratings may be changing over time or flat throughout the period and thus not attributable to GTO support. However, other data support this interpretation. Interviews with site staff (not reported, available upon request) showed that knowledge and performance of GTO steps such as planning processes increased from year 1 to year 2 of implementation for intervention sites but not control sites. Thus, it is likely that CFIR constructs targeted by the GTO implementation support would demonstrate higher ratings over time, while other CFIR constructs not targeted by GTO would remain unchanged. Finally, not all construct coders were blinded to study condition. But one coder for each site was blinded, and that served as a check because the two-person team had to reach consensus for each rating.

Conclusions

Community-based organizations can have difficulty implementing evidence-based programs with fidelity to achieve intended outcomes. This study suggests that strong planning, evaluation, and reflection, likely improved with GTO support, can lead to better program model adherence, a key component of program fidelity. These improvements were demonstrated even in the face of a perception of less desirable implementation readiness including staffing instability and

low resources. These findings highlight that implementation support such as GTO is likely to help lower-resourced community-based organizations improve fidelity through a focus on planning and evaluation processes.

Acknowledgments We would like to acknowledge the contributions of Keisha McDonald, Elizabeth D'Amico, Laura Damschroder, and Christian Lopez for their assistance.

Funding All the authors are funded by a grant from the National Institute on Alcohol Abuse and Alcoholism: Preparing to Run Effective Prevention (R01AA022353-01) to conduct the reported research. The funder had no role in the design of the study, data collection, analysis, interpretation of data, or writing of the manuscript.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with ethical standards of the institutional research committee (RAND Human Subjects Protection Committee FWA00003425) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants in the study.

References

Aarons, G. A., Hoagwood, K., Landsverk, J., Glisson, C., Kelleher, K., & Cafri, G. (2010). Psychometric properties and US national norms of the Evidence-Based Practice Attitude Scale (EBPAS). *Psychological Assessment, 22*(2), 356–365. https://doi.org/10.1037/a0019188.

Acosta, J., Chinman, M., Ebener, P., Malone, P. S., Paddock, S., Phillips, A., et al. (2013). An intervention to improve program implementation: Findings from a two-year cluster randomized trial of Assets-Getting To Outcomes. *Implementation Science, 8*, 87.

Agresti, A. (2002). Inference for contingency tables. *Categorical data analysis*, 70–114.

Benjamini, Y., & Hochberg, B. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, 57*(1), 289–300.

CFIR Research Team (undated). Consolidated Framework for Implementation Research constructs: culture. Center for Clinical Management Research. https://cfirguide.org/constructs/culture/. Accessed 16 July 2019.

Chinman, M., Imm, P., & Wandersman, A. (2004). *Getting to Outcomes 2004: Promoting accountability through methods and tools for planning, implementation, and evaluation*. Santa Monica, CA: RAND Corporation.

Chinman, M., Hunter, S. B., Ebener, P., Paddock, S., Stillman, L., Imm, P., et al. (2008). The Getting To Outcomes demonstration and evaluation: An illustration of the Prevention Support System. *American Journal of Community Psychology, 41*, 206–224.

Chinman, M., Acosta, J., Ebener, P., Malone, P. S., & Slaughter, M. E. (2016). Can implementation support help community-based settings better deliver evidence-based sexual health promotion programs? A randomized trial of Getting To Outcomes(R). *Implementation Science, 11*(1), 78.

Chinman, M., Acosta, J., Ebener, P., Malone, P. S., & Slaughter, M. E. (2018a). A cluster-randomized trial of Getting To Outcomes' impact on sexual health outcomes in community-based settings. *Prevention Science, 19*(4), 437–448. https://doi.org/10.1007/s11121-017-0845-6.

Chinman, M., Ebener, P., Malone, P. S., Cannon, J., D'Amico, E. J., & Acosta, J. (2018b). Testing implementation support for evidence-based programs in community settings: A replication cluster-randomized trial of Getting To Outcomes(R). *Implementation Science, 13*(1), 131. https://doi.org/10.1186/s13012-018-0825-7.

Creswell, J. W., Klassen, A. C., Plano Clark, V. L., & Smith, K. C. (2011). *Best practices for mixed methods research in the health sciences* (pp. 541–545). Bethesda (Maryland): National Institutes of Health.

D'Amico, E. J., & Edelen, M. (2007). Pilot test of Project CHOICE: A voluntary after school intervention for middle school youth. *Psychology of Addictive Behaviors, 21*(4), 592–598.

D'Amico, E. J., Ellickson, P. L., Wagner, E. F., Turrisi, R., Fromme, K., Ghosh-Dastidar, B., et al. (2005). Developmental considerations for substance use interventions from middle school through college. *Alcoholism: Clinical and Experimental Research, 29*, 474–483.

D'Amico, E. J., Tucker, J. S., Miles, J. N. V., Zhou, A. J., Shih, R. A., & Green, H. D. J. (2012). Preventing alcohol use with a voluntary after school program for middle school students: Results from a cluster randomized controlled trial of CHOICE. *Prevention Science, 13*, 415–425.

Damschroder, L. J., & Lowery, J. C. (2013). Evaluation of a large-scale weight management program using the Consolidated Framework for Implementation Research (CFIR). *Implementation Science, 8*, 51. https://doi.org/10.1186/1748-5908-8-51.

Damschroder, L. J., Aron, D. C., Keith, R. E., Kirsh, S. R., Alexander, J. A., & Lowery, J. C. (2009). Fostering implementation of health services research findings into practice: A consolidated framework for advancing implementation science. *Implementation Science, 4*, 50–65. https://doi.org/10.1186/1748-5908-4-50.

Damschroder, L. J., Reardon, C. M., Sperber, N., Robinson, C. H., Fickel, J. J., & Oddone, E. Z. (2017). Implementation evaluation of the Telephone Lifestyle Coaching (TLC) program: Organizational factors associated with successful implementation. *Translational Behavioral Medicine, 7*(2), 233–241. https://doi.org/10.1007/s13142-016-0424-6.

Dedoose Version 8.0.42 (2016). Web application for managing, analyzing, and presenting qualitative and mixed method research data www.dedoose.com.

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*, 327–350.

Ennett, S. T., Ringwalt, C. L., Thorne, J., Rohrbach, L. A., Vincus, A., Simons-Rudolph, A., et al. (2003). A comparison of current practice in school-based substance use prevention programs with meta-analysis findings. *Prevention Science, 4*(1), 1–14.

Glasgow, R., Vinson, C., Chambers, D., Khoury, M., Kaplan, R., & Hunter, C. (2012). National Institutes of Health approaches to dissemination and implementation science: Current and future directions. *American Journal of Public Health, 102*(7), 1274–1281.

Hallfors, D., & Godette, D. (2002). Will the 'principles of effectiveness' improve prevention practice? Early findings from a diffusion study. *Health Education Research, 17*(4), 461–470.

Hawkins, J. D., Oesterle, S., Brown, E. C., Arthur, M. W., Abbott, R. D., Fagan, A. A., et al. (2009). Results of a type 2 translational research trial to prevent adolescent drug use and delinquency: A test of Communities That Care. *Archives of Pediatrics and Adolescent Medicine, 163*(9), 789–798.

Helfrich, C. D., Kohn, M. J., Stapleton, A., Allen, C. L., Hammerback, K. E., Chan, K. C. G., et al. (2018). Readiness to change over time: Change commitment and change efficacy in a workplace health-promotion trial. *Frontiers in Public Health, 6*. https://doi.org/10.3389/fpubh.2018.00110.

Johnston, L. D., O'Malley, P. M., Miech, R. A., Bachman, J. G., & Schulenberg, J. E. (2017). In The University of Michigan (Ed.), *Monitoring the future national survey results on drug use, 1975-2016: Overview, key findings on adolescent drug use*. Ann Arbor, Michigan: Institute for Social Research.

Kirk, M. A., Kelley, C., Yankey, N., Birken, S. A., Abadie, B., & Damschroder, L. (2016). A systematic review of the use of the Consolidated Framework for Implementation Research. *Implementation Science, 11*(1), 72. https://doi.org/10.1186/s13012-016-0437-z.

Kumar, R., O'Malley, P. M., Johnston, L. D., & Laetz, V. B. (2013). Alcohol, tobacco, and other drug use prevention programs in U.S. schools: A descriptive summary. *Prevention Science, 14*(6), 581–592. https://doi.org/10.1007/s11121-012-0340-z.

Livet, M., & Wandersman, A. (2005). Organizational functioning: Facilitating effective interventions and increasing the odds of programming success. In D. M. Fetterman & A. Wandersman (Eds.), *Empowerment evaluation in practice* (pp. 123–154). New York, NY: Guilford.

Miller, W. R., & Rollnick, S. (2012). *Motivational interviewing: Helping people change* (3rd ed.). New York: Guilford Press.

Ringwalt, C., Hanley, S., Vincus, A. A., Ennett, S. T., Rohrbach, L. A., & Bowling, J. M. (2008). The prevalence of effective substance use prevention curricula in the nation's high schools. *Journal of Primary Prevention, 29*(6), 479–488. https://doi.org/10.1007/s10935-008-0158-4.

Ringwalt, C., Vincus, A. A., Hanley, S., Ennett, S. T., Bowling, J. M., & Rohrbach, L. A. (2009). The prevalence of evidence-based drug use prevention curricula in U.S. middle schools in 2005. *Prevention Science, 10*(1), 33–40. https://doi.org/10.1007/s11121-008-0112-y.

Scaccia, J. P., Cook, B. S., Lamont, A., Wandersman, A., Castellow, J., Katz, J., et al. (2015). A practical implementation science heuristic for organizational readiness: $R=MC^2$. *Journal of Community Psychology, 43*(4), 484–501.

Spoth, R., Redmond, C., Shin, C., Greenberg, M., Clair, S., & Feinberg, M. (2007). Substance-use outcomes at 18 months past baseline: The PROSPER Community-University Partnership Trial. *American Journal of Preventive Medicine, 32*(5), 395–402.

U.S. Department of Health and Human Services. (2016). *Facing addiction in America: The Surgeon General's report on alcohol, drugs, and health* (Vol. 6). Washington, DC: HHS.

Weiner, B. J. (2009). A theory of organizational readiness for change. *Implementation Science, 4*(1), 67. https://doi.org/10.1186/1748-5908-4-67.