American Journal of Evaluation
http://aje.sagepub.com

Is Sustainability Possible? A Review and Commentary on Empirical Studies of Program Sustainability

Mary Ann Scheirer

American Journal of Evaluation 2005; 26; 320
DOI: 10.1177/1098214005278752

The online version of this article can be found at:
http://aje.sagepub.com/cgi/content/abstract/26/3/320

Published by:
SAGE Publications
http://www.sagepublications.com

On behalf of:
American Evaluation Association

Is Sustainability Possible?

A Review and Commentary on Empirical Studies of Program Sustainability

Mary Ann Scheirer
Scheirer Consulting

Abstract: An important final step in the life cycles of programs and their evaluation involves assessing new programs' or innovations' sustainability. This review and synthesis of 19 empirical studies of the sustainability of American and Canadian health-related programs examines the extent of sustainability achieved and summarizes factors contributing to greater sustainability. Three definitions for measuring sustainability were examined: continued program activities (18 studies), continued measured benefits or outcomes for new clients (2 studies), and maintained community capacity (6 studies). Methods of studying sustainability were also assessed. In 14 of 17 studies covering the continuation of program activities, at least 60% of sites reported sustaining at least one program component. Although these studies' methods had substantial limitations, cross-study analysis showed consistent support for five important factors influencing the extent of sustainability: (a) A program can be modified over time, (b) a "champion" is present, (c) a program "fits" with its organization's mission and procedures, (d) benefits to staff members and/or clients are readily perceived, and (e) stakeholders in other organizations provide support.

Keywords: *sustainability; health promotion programs; routinization; institutionalization; program life cycle; evaluation*

The topic of sustainability is increasingly important to the funders and implementers of health-related demonstration programs and innovations. What happens after the initial funding for new programs expires? Do the programs continue or end their activities or even expand to new sites or new beneficiaries? Does the concept of "seed funding" have validity in encouraging the start-up of new programs that are then continued by other means? In health-related content fields, several sources of major funding exist for "demonstration" projects, including federal government agencies and foundations. These projects typically receive funding for a few years—usually only 3 to 5 years—and then are expected to obtain other funding and resources for continuation.

For example, at the Robert Wood Johnson Foundation, at which I worked at the time this review was initiated, many staff members questioned whether the projects funded within its tar-

---

Mary Ann Scheirer, Scheirer Consulting, 238 Sayre Drive, Princeton, NJ 08540; e-mail: maryann@scheirerconsulting.com.

Author's Note: I want to acknowledge the invaluable help of Kathryn Flatley of the Robert Wood Johnson Foundation, who provided the literature search on which this review is based, and to express appreciation for the many useful suggestions on earlier drafts of the article by Laura Leviton, Amardeep Thind, and the editors and anonymous reviewers of the *American Journal of Evaluation.*

American Journal of Evaluation, Vol. 26 No. 3, September 2005 320-347
DOI: 10.1177/1098214005278752

geted initiatives would survive after relatively short periods of external funding. Some staff members were rather pessimistic about the chances for the survival of these projects, given the financial stresses facing most publicly funded programs. I was working with the evaluations of nearly 20 projects for addressing pediatric asthma, which were facing their final years of funding and trying to plan for sustainability. What prior studies could provide evidence to guide them about the extent to which sustainability might be possible and under what conditions? I initiated this review to find out what empirical evidence was available on this topic in the health field. However, the topic of sustainability is likely to be relevant to many other content fields, particularly those using short-term grants as funding sources. Evaluators in other fields might benefit from the findings and discussion in this review about the methods used, the tentative findings, and the suggestions for improved future evaluation of program sustainability in diverse content fields.

This article reports the results of a systematic review of empirical literature on the sustainability of health-related projects, focusing on studies that report data collected at a time point after the initial external funding had expired, for programs or innovations related to health or health care. Studies on this topic were found in several different categories, including sustainability, institutionalization, maintenance, durability, and continuation. The purpose here is to examine the types and extent of sustainability achieved for the programs studied, as well as to summarize findings concerning factors that were found by researchers to contribute to greater likelihood of sustainability.

A program is defined here as a set of resources and activities directed toward one or more common goals (Newcomer, Hatry, & Wholey, 1994). In many cases, the federal or foundation programs that were the subject of the studies reviewed here had funded sets of local projects for limited time periods. These projects were intended to show benefits for some group of clients in relation to a health topic, including heart health interventions, screening for breast cancer, and support for frail elderly, among other intervention topics. Most of the studies reviewed here assessed the continuation of multiple local projects at their original sites, using an organizational unit of analysis.[1]

I address here neither the continued funding of federal- or foundation-sponsored programs as a whole nor the potential transfer of projects to other sites or dissemination to new sites. In addition, I do not include here the sustainability of organizations per se, which is another important issue that may influence the sustainability of projects funded within those organizations. Several other articles discuss developing the capacity of entire communities to sustain programs, for example, focusing on the relationships between researchers who develop community health interventions and the capacity of communities to continue these interventions (Altman, 1995) or on how to build "capacity" in communities for health promotion (Hawe, Noort, King, & Jordens, 1997; Labonte & Laverack, 2001). But those articles did not report empirical evidence on the extent of sustainability in communities after specified interventions, so they are not included in the reporting of empirical studies in this article. Furthermore, studies of the maintenance of desired behaviors among individual clients (compared with relapse), such as continued abstinence from smoking or maintaining sobriety after substance abuse, were not included in this review. These alternative perspectives on "sustainability" and other types of positive outcomes from the programs are also very important topics, and each would need substantial review work to address appropriately.

The growing literature on the general theme of what happens to projects after their initial funding ends has not yet coalesced into a single research paradigm, a shared set of statistical methods, or even a common terminology. Therefore, this review of findings included a broad set of studies to explore what general findings, if any, could be gleaned from them. I did not attempt the formal statistical methods of meta-analysis, because most of the original studies did

not report the statistically derived findings needed for meta-analysis. However, a seminal literature review and theoretical framework by Shediac-Rizkallah and Bone (1998) provides important theoretical underpinnings and categories for this review, even though many authors of the studies reviewed did not use that framework (or wrote prior to Shediac-Rizkallah and Bone's review).

Context

The topic of sustainability is often set within a life cycle perspective about program development, implementation, evaluation, maintenance, and sometimes dissemination to other sites or beneficiaries (Livit & Wandersman, 2004; Pluye, Potvin, & Denis, 2004; Scheirer, 1990; Wandersman, Imm, Chinman, & Kaftarian, 2000; Yin, 1981). From this perspective, new programs intended to improve health or other services often proceed through a series of overlapping stages, such as the following:

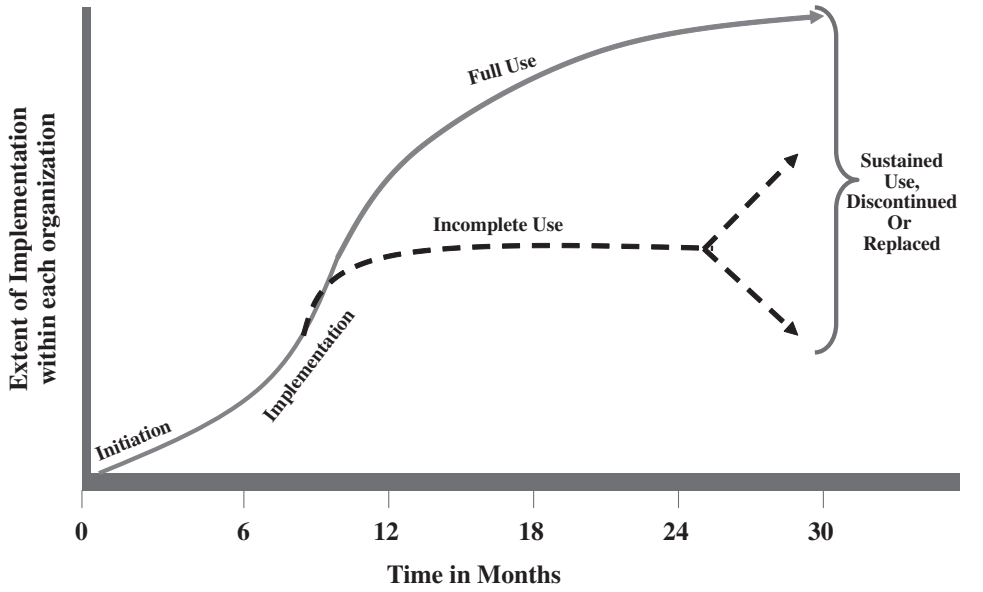
- Initiation: A program idea is conceived by innovators within an organization or by researchers wanting to test a potential new solution to a problem or felt need.
- Development and adoption: The program idea is "fleshed out" with details about its components and activities, to be tested in the intended real-life context. Alternatively, a program developed in a different location is adopted by the target organization or community.
- Implementation: The program ideas are put into full practice within the target organization or community; implementation often may require a year or more of work to define work roles, train staff members, acquire necessary technology or other resources, and try out delivery options, as well as to secure needed administrative and/or community support.
- Sustainability (or discontinuation): The program components developed and implemented in earlier stages are (or are not) maintained after the initial funding or other impetus is removed.
- Dissemination: For some programs, the funding organization expects the new program idea to be communicated to other sites and adapted or replicated to serve new beneficiaries.

This model is portrayed in Figure 1, showing the various stages as taking place over a considerable period of time, often several years. The time frames for these stages shown in Figure 1 are illustrative only, not based on empirical data about how long each stage might require. Frequently, the processes in each stage are overlapping. Yet the fact that grants for such projects are often awarded for a time period of 3 years or shorter suggests that funders expect the program development period to follow a trajectory similar to that illustrated in Figure 1 to achieve measurable outcomes for clients by the end of the grant period. This projected time schedule may be overly optimistic for many new programs.

Evaluation is not presented as a specific stage, because different types of evaluation should accompany each stage. For example, formative evaluation is used to illuminate the development period, then process and outcome evaluation track the delivery and results of the implementation stage. Dissemination would require additional evaluative data collection to track the extent and results of use by additional sites. Assessing sustainability requires further data collection to examine whether the activities and benefits of the implementation phase continue. This stage is the topic of this article, focusing on an organizational unit of analysis. I do not include here studies that assessed the dissemination, adaptation, or replication of the initial program model, except for a few studies that consider the extent of sustainability within new sites.

Although these analytical stages can be identified as a theoretical framework, in actual sites, they are often overlapping and sometimes nonlinear. That is, some implementation occurs during the stage of developing the detailed components, particularly for a new program idea. Dissemination to other sites may occur before the program is fully implemented or sustained in the

Figure 1
Program Life Cycle



original site. And, as shown in Figure 1, implementation is often incomplete or at a limited status of program delivery before the initial funding terminates. In this case, sustainability may be difficult, because the intended program activities were never fully carried out. A recent article by Pluye et al. (2004) even questions the viability of the concept of stages, proposing that implementation and sustainability are parallel processes that occur concomitantly.

A key point for this setting of context is that sustainability is likely to be affected by all the preceding program activities, as well as by its funding arrangements. For example, a program that is generated at the initiation stage from a strong internal need felt by those within an organization may be more likely to be sustained than one generated from an external "push," such as the availability of funding from an outside agency. Furthermore, the same organizational characteristics that foster strong implementation of a new program, such as its compatibility with the organization's mission and the involvement of strong support by organizational "champions," are likely to enable continued delivery as sustainability.

Conceptual Frameworks

Research on the general topic of "what happens after the funding ends" for a specific program is not yet well conceptualized into agreed-on methods and topics. Various authors tend to approach the topic in very diverse ways. The most inclusive framework for summarizing the available empirical studies is the work by Shediac-Rizkallah and Bone (1998), which reviewed literature available to that date about health intervention programs, both in the United States and internationally. Recent frameworks with similar components have been suggested by Johnson, Hays, Center, and Daley (2004) for interventions in the substance abuse prevention field and by

Mancini and Marek (2004) for family support programs.[2] After examining various definitions for what constitutes program sustainability, Shediac-Rizkallah and Bone suggested that indicators of sustainability fall into three distinct types of measures (i.e., different operational definitions):

1. measuring continued health benefits for individuals after the initial program funding ends, particularly continuing to achieve beneficial outcomes among new consumers or other intended recipients (in contrast to maintaining behavioral change among earlier clients);
2. inquiries concerning the continuation of program activities within an organization, often termed "institutionalization" or "routinization," within an organizational focus; and
3. questions about the continued capacity of a community to develop and deliver health promotion programs, particularly relevant when the initial program worked via a community coalition or other community capacity–developing process.

These three types of measures relate to different units of analysis for sustainability measurement: individual-level outcomes, the organization-level implementation of activities, and community-level capacity. Conceptually, these three levels might reasonably relate to linked components of a program's logic model (community capacity → sustained program activities → outcomes for clients). Yet in the literature to date, these three types of sustainability measures are usually addressed separately and are not interchangeable. For example, a community coalition can continue to meet and may plan for relevant activities without necessarily implementing programs that provide benefits for intended clients, the purpose that generated the formation of the coalition. A program might continue in existence and implement some activities but not document measurable outcomes for its clients.[3] Furthermore, the routinization of a program within an ongoing organization could result in the indefinite continuation of its client outcomes or could become a hollow shell of activities perpetuated for their own sake, whether or not benefits for clients are achieved. At least one observer has questioned whether institutionalization ought to be a goal of the life cycle of program development and delivery (Green, 1989). Green (1989) suggested instead that the capacity building and innovativeness generated by the development of new programs is the more important outcome that should be sustained. A fourth type of sustainability is suggested by work of the Harvard Family Research project: sustaining the ideas, beliefs, principles, or values underlying an initiative (Weiss, Coffman, & Bohan-Baker, 2002). However, this focus on the cognitive components of programs was not addressed by the literature reviewed here and might be difficult to operationalize for systematic research.

This review does not encompass the topic of organizational sustainability, that is, whether funded organizations, especially nonprofit organizations, are maintained over long periods of time. This topic tends to be addressed in the separate literatures about organizational behavior, organizational development, and entrepreneurship and was not included in this review. Funders who support the start-up of new organizations to house new programs might well consider both the sustainability of the program activities and the sustainability of the organization (cf. Livit & Wandersman, 2004).

Influences on Sustainability

The previous literature about program sustainability has also investigated the question, What factors help increase the likelihood of sustainability? This question is particularly addressed in the literature concerning the institutionalization of program activities within an organization. This issue is of central importance when one is planning for program sustainability, when it is helpful to know what processes and other influences need to be considered to extend the deliv-

ery of program activities. Although program staff members often think first about finding new sources of funding for a program whose initial funding will end soon, the influences suggested by the literature extend far beyond simply bringing in new funding. Shediac-Rizkallah and Bone (1998) provided a useful framework of potentially influential factors, which I adapted for this review. The framework includes the following components:

- *Aspects of project design and characteristics:* These include the nature of the start-up and design process, particularly whether local stakeholders were involved; whether the program is modifiable to meet local needs and conditions; whether evaluation has documented effectiveness; how long the program or innovation has existed; and the nature of the original sources of financing.
- *Factors within the organizational setting:* These include whether there is a program champion who is strategically placed to foster continuation, whether the new program or innovation is congruent with the underlying mission and operating procedures of the organization, and the underlying capacity of the organization (sometimes shown by its length of time in existence). The work of Robert Yin (1979, 1981) on factors associated with routinization (see below) amplifies this part of the framework.
- *Factors in the broader community environment:* These include the stability and favorability of external socioeconomic and political factors, such as market forces impinging on an organization, legislation affecting the program, support from external community leaders, and the availability of funding and other resources as inputs to the program.

These influences on the extent of program sustainability are not discrete variables whose strength of effects can be easily tested in isolation from one another. Instead, they are likely to interact over time to weave the history of each program or innovation. A factor that was crucially important to the longevity of one program may have been unimportant in the "story" of another site's implementation of the same program. Furthermore, few of the sources reviewed for this article considered the same set of potential influences on the observed extent of sustainability or operationalized them in the same ways. For these reasons, research on the topic of program sustainability, although greatly needed, is not likely to develop and validate a single set of guidance about "how to do it." Similar to the results of research on program implementation (Scheirer, 1981, 1987), research about program sustainability is likely to remain multifaceted, with results contingent on the specific programs and contexts in which they are operating. Yet future research that builds on the methods and findings of the studies reviewed here is strongly needed to consolidate empirical evidence and to test strategies aimed at increasing the numbers of sustained programs from the moderate levels reported below.

Sustainability as Institutionalization

One detailed line of research within the broader topic of program sustainability defines sustainability as the institutionalization or routinization of programs into ongoing organizational systems. In this perspective, the maintenance of program activities without special external funding is most likely to occur if the program components become embedded into organizational processes. When this happens, researchers may no longer be able to identify a specific "program," because the program activities have become a part of the organization's core services. These concepts are well developed in Yin's (1979, 19891) concept of routinization. Using the results of 19 case studies of technical innovations introduced into local governments in the 1970s (such as police computer systems, mobile intensive care units for paramedics, and alcohol Breathalyzer testing for driver safety), Yin examined how these innovations became part of standard practice. He suggested that full routinization depends on 12 processes or events that he characterized as specific "passages" and "cycles," listed in Table 1. Using these catego-

Table 1
Yin's (1979, 1981) Routinization Framework

1. Budget	1a. Program supported by change from soft to hard money 1b. Survives annual budget cycles
2. Personnel	2a. Program activities become part of job descriptions/requirements 2b. Program survives turnover of personnel/leadership 2c. Key program staff members are promoted within agency 2d. Program activities spread to all potential users within agency
3. Supply and maintenance	3a. Supply and maintenance provided by agency 3b. Activities survive equipment turnover
4. Training	4a. Skills taught in many training cycles 4b. Skills become part of professional standards
5. Organizational governance	5a. Use of program recognized in manuals, procedures, regulations 5b. Program recognized as permanent within agency

ries, he distinguished three degrees of routinization: marginal, moderate, or high, reflecting the number of the passages and cycles that had been achieved. His analysis found that important conditions for routinization were internal to the specific local agency, particularly the extent to which it was supported by local staff members using it and the support of top agency administrators (Yin, 1981).

The work of Goodman, McLeroy, Steckler, and Hoyle (1993) extended Yin's (1979, 1981) framework by developing a questionnaire measuring tool for the Level of Institutionalization. They categorized processes similar to Yin's into four organizational subsystems (production, maintenance, supportive, and managerial). The instrument further extends the scope of program delivery measurement by adding "niche saturation," the extent to which each component is fully embedded into all relevant subsystems. The article presenting this instrument proposed an eight-factor model, with supporting data from administrators in 141 organizations. However, this model has been questioned as not fully supported by the data (Scheirer, 1993), and further research using it (Barab, Redman, & Froman 1998) suggests that a two-factor model (labeled the presence of "routines" and "niche saturation") provides improved reliability and validity. To date, only one later study of sustainability (Goodson, Smith, Evans, Meyer, & Gottlieb, 2001) was located that used the Level of Institutionalization scale, and it was a version adapted for a qualitative, case-study methodology.

These issues about the conceptualization and measurement of program sustainability, as well as the conceptual frameworks for assessing factors which are likely to influence the extent of sustainability, form the background for this review of empirical literature. I now turn to the methods and findings of the literature review itself.

Methods for This Review

For this literature review, the first step was to search for research that had collected data about some aspect of program sustainability after initial funding had ended. The search was limited to studies in the health arena in the United States and Canada and did not include other studies about program sustainability conducted in the international arena. The health field has used the "demonstration model" extensively, for which the question of postfunding sustainability has become an important issue. I expected that there would be greater convergence among studies within the same field than might occur across multiple content areas or

among diverse environments internationally. Furthermore, this review was initiated under the auspices of the Robert Wood Johnson Foundation, which focuses only on health and health care goals. Whether the findings would be replicated for projects in other content fields is an important question needing further investigation.

The search was conducted using the search string "sustainability OR routinization OR institutionalization AND health OR healthcare," in all major relevant bibliographic databases, for the years 1990 to 2003, including PubMed, ProQuest, the Librarians Index to the Internet, and NLM Gateway. The abstracts of potentially relevant citations were examined to determine if the original research included data collected about any aspect of sustainability after the initial funding had ended. Full texts of all relevant articles were then obtained. A few studies were already known to me from prior related work. In addition, reference lists of obtained articles were examined for any additional studies, such as those using different terminology. The systematic review did not include articles or how-to-do-it commentaries about sustainability that did not report empirical data, although these articles were consulted for their conceptual frameworks and approaches. These procedures yielded 19 studies that met the criteria for inclusion: reporting data collected about the status and/or influences on health program sustainability (including case studies). The review included all available studies that met these criteria, not a sample of them.

The next step was to code information from each study into a set of tables, so that overall results could be tabulated and compared. I did the coding using the categories suggested in the framework developed by Shediac-Rizkallah and Bone (1998). Extracts from the article coding appear in Tables 2 to 4. As much as possible, I extracted factual material from each original article. However, in many cases, the original articles used narrative description about the likely influences on sustainability rather than a set of variables with explicit definitions. Therefore, the components from each study coded as showing an influence on sustainability, in Table 4, required some subjective judgment.

Findings

Methods for Studying Sustainability

Using the search strategies listed above, 19 studies were located and analyzed for this article. A first analytic set of questions concerns the methods used to study sustainability: What methods were used, and how systematic were the methods used to derive the findings of these studies? Table 2 shows the scope of each study (the number of sites or programs studied) as well as several features of its methods.

Numbers and selection of sites. The number of sites or programs studied ranged from 5 to 787 sites in one large program. Seven studies examined fewer than 20 sites, 5 studies looked at 20 to 50 sites or programs, and 7 analyzed more than 50 sites or programs. The selection of sites and interventions within each article varied: Many attempted to contact all the sites funded by the programs they were assessing or identified multiple interventions within targeted locations. Two articles selected sites known to have survived or not (Glaser, 1981) or to show a range of sustainability (Goodman & Steckler, 1989); these articles were not included in the tabulation of the extent of sustainability for this review. This review uses the 19 studies as its unit of analysis (not the sites within individual studies). There is no way to know whether the findings from these studies would apply to some larger population of programs, because none of the studies

(text continues on p. 334)

Table 2
Methods Used in Sustainability Research

Author and Date	Number of Programs/Sites	Methods of Data Collection	Number of Respondents per Site	Time Since Funding Ended	Any Statistical Tests?	Any Data Quality Checks?
Bracht et al. (1994)	78 programs at 3 sites	? Mail survey, key informants	Multiple	3 years	No	?
Elder et al. (1998)	22 interventions at 1 site	? "Monitoring"	Not specified	> 2 years	No	No
Evashwick and Ory (2003)	20 projects: won award for "best practice models"	Telephone? interviews	One per project	Not stated: varied funding status	No	Some
Glaser (1981)	18 programs	Site visits after telephone contacts	Multiple?	Survived > 2 years	Yes	Yes
Goodman and Steckler (1989)	10 sites	Case studies using interviews, observation, document review	Multiple	? All in existence from 3 to 6 years	No	Yes
Goodson, Smith, Evans, Meyer, & Gottlieb (2001)	5 primary care sites	Case studies using in-person and phone interviews plus written questionnaire	Multiple	3 years	No	Some
Harris et al. (2003)	7 states	Telephone interviews	Multiple	2 years	No	No
Herrera (2002)	787 local coalitions for Faith in Action program (of 1,091 funded sites)	Mail and telephone survey, documents, early project reports	One per site (?)	Varied: up to 6 years	Yes	Some
Hogg, Baskerville, Nykiforuk, & Mallen (2002)	7 family medical practices, Canada	Group interviews (doctors, nurses, staff members)	Multiple	No external funding; 1 year after intervention	No	Yes
Lichtenstein, Thompson, Nettekoven, & Corbett (1996)	11 communities	Site visit, focus groups; phone interviews	Multiple	Approximately 1 year	No	Some
Lodl and Stevens (2002)	56 community coalitions: rural	Telephone interviews (regarding 18 coalitions)	Not specified	5 years	No	No
O'Loughlin, Renand, Richard, Sanchez-Gomez, & Paradis (1998)	189 interventions in 30 sites	Telephone interviews	One per intervention	0 to > 20 years	Yes	Yes
Paine-Andrews, Fisher, Campuzano, Fawcett, & Berkley-Patton (2000)	6 community initiatives	Mailed survey (4 of 6 sites); interviews	Not stated	1 year	No	No

Scheirer (1990)	490 school districts	Telephone interviews	1 interview per district	Varied: mean of 6 years	Yes	Some
Shediac-Rizkallah, Scheirer, & Cassady (1997)	28 hospitals	Telephone interviews; surveillance database	One per hospital, plus mammogram records	6 months	Yes	Yes
Stevens and Peikes (2004)	112 local health projects (of 120 funded)	Telephone survey plus 10 case studies	Multiple	Varied, average of 4.8 years	No: too few nonsustained	Some
Stange, Goodwin, Zyzanski, & Dietrich (2003)	37 family medical practices	Data extracted from medical records by researchers	11 to 40 patient records per physician	12 months after intervention	Yes	No?
Thompson, Lichtenstein, Corbett, Nettekoven, & Feng (2000)	11 communities in smoking program, plus 11 comparison communities	Mailed questionnaires to key informants	Multiple	2 years after intervention	Yes	Yes
Wallin, Bostrom, Wikblad, & Ewald (2003)	119 nurses, in Sweden, from 75 organizations	Mailed questionnaire	One: individuals were focus of study	4 years after training	Yes	Yes

Table 3
Types of Sustainability Assessed

Author and Date	Type of Program	Number of Years: External Funding	Operational Definition of Sustainability?	% of Sites Sustained?	Types of Sustainability Examined: Benefits to Clients	Program Activities	Community Support
Bracht et al. (1994)	Heart health interventions	Up to 8 years	Yes	60% of programs	—	Yes	Yes
Elder et al. (1998)	Nutrition and heart health	2 to 3 years	No: narrative only	6 (2 major ones) of 22 interventions (27%)	Examples given, with data	Yes	Briefly
Evashwick and Ory (2003)	Gerontological health: varied	? Projects were from 2 to 15 years; funding source not known	No	At least 5 of 20 "dormant," 75% sustained (?)	Examples of numbers served	Yes	Briefly
Glaser (1981)	Mental health: Goal Attainment Scaling and Fairweather Lodges	Not applicable: not selected on funding status	Sites selected on basis of > 2 years survival vs. not	Selected for case studies: about half sustained	—	Yes	Some
Goodman and Steckler (1989)	Community health promotion	3 years	Yes: used Yin's (1979, 1981) model	40% (1 rated high, 3 moderate institutionalization) Selected for range	—	Yes	—
Goodson, Smith, Evans, Meyer, & Gottlieb (2001)	Put Prevention Into Practice at 5 primary care sites	3 years	Yes: based on Level of Institutionalization scales	4 of 5 (80%) at 3 time periods	Not after funding	Yes	Some
Harris et al. (2003)	Multidisciplinary, community-based education for health professionals	5 years	No: part of interview questions	5 of 7 (71%) "sustained as originally planned"; also 63 policy changes	—	Yes	? mentioned plus
Herrera (2002)	Faith in Action: coalitions of religious and other organizations use volunteers for supporting homebound clients	18 months for most, some received funds for a few additional months	Brief: continuing to serve clients	At least 62% continue to serve clients; perhaps up to 78%	Brief: only numbers of clients	Some, especially regarding support for volunteers	Some
Hogg, Baskerville, Nykiforuk, & Mallen (2002)	External facilitator to increase prevention in primary care practices	None: facilitation only	No: informal	Varied by intervention: 6 of 7 (86%) continued one intervention; others lower percentages	—	Yes	—

Lichtenstein, Thompson, Nettekoven, & Corbett (1996)	Community-based smoking cessation	4 years	Brief	9 of 11 (82%) kept coalition; 68% of initial activities	—	Yes	Yes
Lodl and Stevens (2002)	Community coalitions for children's and families' needs	5 years	No	Only 18 of 56 coalitions could be contacted (32%); 5 (28% of contacted) still "active"	—	Yes	Slight
O'Loughlin, Renaud, Richard, Sanchez-Gomez, & Paradis (1998)	Heart health promotion interventions	Varied: median of 24 months	Yes	44% "very permanent"; 35% "somewhat permanent"	—	Yes	—
Paine-Andrews, Fisher, Campuzano, Fawcett, & Berkley-Patten (2000)	Teen pregnancy and teen substance abuse: in communities	4 years	Yes	100% of 6 communities kept projects; 32% of 378 activities remained		Yes	Yes
Scheirer (1990)	Dental program in public schools	Extent of external funding unknown: schools used program for mean of 6 years	Yes	79% still using program when surveyed	—	Yes	—
Shediac-Rizkallah, Scheirer, & Cassady (1997)	Breast cancer screening program for low–socioeconomic status women	5 years	Yes	44% of number of screening mammograms; 18 of 28 (64%) reported maintaining some program activity	Yes	Yes	Yes
Stevens and Peikes (2004)	Diverse types of local health-related projects	3 to 4 years	Brief	At least 75% of projects still in operation; 92% survived at least 1 year	—	Yes	Some
Stange, Goodwin, Zyzanski, & Dietrich (2003)	Interventions to increase use of prevention services	Not applicable: intervention was in-person assistance, no external funding	Yes	Not analyzed: only overall rate of services: no significant reduction by 12 months after inventions	Yes	—	—
Thompson, Lichtenstein, Corbett, Nettekoven, & Feng (2000)	Smoking cessation: in communities	4 years	Yes	9 of 11 (82%) communities in both intervention and control sites had tobacco control structures; activity levels in 5 intervention areas had similar ratings: moderate to high	—	Yes	Yes
Wallin, Bostrom, Wikblad, & Ewald (2003)	4-day training on quality improvement	Not applicable	Yes: briefly	46 of 119 (39%)	—	Yes	—

Table 4
Findings: Factors Reported to Be Related to the Extent of Sustainability

	Project Design and Characteristics				Organizational Setting					Community and Environment	
Author and Date	% of Sites Sustained?	Program Modifiable	Low Cost; Uses Volunteers	Positive Evaluation	Champion Present	Strong Existing Capacity	"Fit" With Mission/ Tasks	Perceived Benefits to Staff Members/ Clients?	Longer Time Period	Support From Other Organizations	Type of Funding
Bracht et al. (1994)	60%	Yes					Yes	Yes		Yes	
Elder et al. (1998)	27%		Yes		Yes		Yes	Yes			
Evashwick and Ory (2003)	75%	Yes	Yes	Yes	Yes	Yes	Yes			Yes	Yes
Glaser (1981)	Not applicable[a]	Yes		Prior	Yes		Yes	Yes		Yes?	Maybe
Goodman and Steckler (1989)	40%[a]	Yes			Yes	Yes	Yes	Yes			
Goodson, Smith, Evans, Meyer, & Gottlieb (2001)	80%				Yes	Yes	Yes	Yes?		Yes	
Harris et al. (2003)	71%				Yes	Negative	Yes			Yes	Yes
Herrera (2002)	62% to 78%		Yes		Yes?				NR	Yes	Yes
Hogg, Baskerville, Nykiforuk, & Mallen (2002)	86%	Yes			Yes?		Yes	Yes			
Lichtenstein, Thompson, Nettekoven, & Corbett (1996)	82% coalitions; 68% activities			Results not available				Yes		Yes	
Lodl and Stevens (2002)	28%	Yes						Yes	Yes	Yes	Yes
O'Loughlin, Renaud, Richard, Sanchez-Gomez, & Paradis (1998)	79%	Yes	Yes	Yes	Yes	NR	Yes	Yes	NR		NR
Paine-Andrews, Fisher, Campuzano, Fawcett, & Berkley-Patton (2000)	100% communities; 32% activities	Yes	Yes		Yes		Yes	Yes		Yes	Yes
Scheirer (1990)	79%				Yes			Yes	Yes	Yes	Yes

Shediac-Rizkallah, Scheirer, & Cassady (1997)	64% hospitals; 44% mammo-grams	Yes			Yes		Yes			Yes	Yes
Stevens and Peikes (2004)	75%	Yes		About half	Yes	Yes	Yes	Yes	Yes?	Yes	Yes
Stange, Goodwin, Zyzanski, & Dietrich (2003)	Not analyzed	Yes		Feedback provided							
Thompson, Lichtenstein, Corbett, Nettekoven, & Feng (2000)	82% of both intervention and control sites	Yes									
Wallin, Bostrom, Wikblad, & Ewald (2003)	39%										
Totals: number of studies citing each factor		12	5	4	13	4	12	12	3 related, 2 NR	12	9, 2 NR

NOTE: Influences are those cited by authors of original studies, not necessarily by statistical tests. NR = factor examined by study but found to be not related to extent of sustainability. Prior = prior evidence of effectiveness available from other research.

a. Projects were selected to show a range of outcome levels on sustainability.

began by identifying any such population. Furthermore, the discussion below about the influences on sustainability also depends on the extent of variability in these "predictors" among the sites studied: If a hypothesized predictor does not vary among sites, then it would not show a statistical association with the dependent variable, sustainability.

Data collection. Not surprisingly, the major methods used to gather data also varied substantially, with mail (5 studies) or telephone (11 studies) surveys being most prevalent. Case-study methods or site visits were used in 6 studies. Other methods or multiple methods were used in 8 studies. One article did not specify the study's methods, merely stating that data were collected by "monitoring." Even though most programs took place in organizations or communities, 6 of the studies used only one informant per site to report on its current status,[4] and 3 others did not specify how many respondents were contacted from each site. More positively, 9 studies contacted multiple respondents per site or program to obtain a more reliable base of information about the project, and 1 study extracted data from medical records about prevention services provided to individuals. About two thirds of the studies (13) reported at least some information about data quality checks or methods undertaken to ensure quality (such as steps to ensure adequate response rates in a survey, methods to assess the reliability or validity of data, or comparisons across different types of data collection to provide triangulation of findings).

Timing of sustainability measurement. Most studies waited for a meaningful amount of time to elapse before examining sustainability: At least 13 studies contacted the original sites at least 2 years after external funding had ended (if there was external funding) or had a varied length of time after funding ended before data collection.[5] There is no commonly accepted time point for defining when a program is "sustained." These studies typically contacted the sites studied at only one time point within a long trajectory of events that might affect sustainability. For some interventions, institutionalization within an ongoing organization might take place rather quickly, if the organizational factors were favorable (such as those identified by Yin, 1981, discussed above). On the other hand, the pace of organizational change is often slow; there might be a tendency to keep staff members on the payroll for a time, to maintain only some activities of a broader initiative, or to keep a recent initiative going for political or face-saving reasons, even if it is not sustained permanently. Detailed investigation about the long-term processes of sustainability or institutionalization of the targeted interventions over a period of several years was generally not undertaken by the studies reviewed.

Statistical analysis. Nearly half of the studies (8 of 19) used some type of statistical analysis or tests of significance for assessing influences related to sustainability, whereas the other 11 reported only narrative data or frequencies of cases. However, the statistical analysis used was often only a bivariate test of these associations, not a more rigorous multivariate analysis to control for the correlations among the influences (e.g., in O'Loughlin, Renaud, Richard, Sanchez-Gomez, & Paradis, 1998; Scheirer, 1990).[6] The relative scarcity of tests of statistical association made it impossible to rigorously compare influences on sustainability across studies, such as by using statistical meta-analysis methods.

This overall assessment of the methods used in these studies presents a mixed picture. Although a few studies were quite rigorous and explicit in the methods they used, several others did not report much information about how they arrived at their conclusions. Some mentioned that funding was not adequate to use more systematic methods. The great majority used only self-report surveys of key staff members from the target projects, usually the project director, and had no data source to cross-validate the responses provided. Only a few studies provided explicit operational definitions of what was meant by "sustainability," and even fewer

operationalized potential influences on this outcome (sustainability) in ways that could be compared across studies. The relative absence of information about the quality of the data underlying findings presented is a substantial weakness in this body of work. Many authors were eager to provide advice about how to increase sustainability but did not always ground this advice on a methodologically sound evidence base. The findings from these 19 studies, reported below, are thus based on a body of rather weak evidence; future studies with more rigorous methods might reach quite different conclusions.

Extent and Types of Sustainability

The types of programs that were examined in these studies varied considerably, as shown in Table 3, so the nature and meaning of sustainability also varied by context. Several studies were of community-based coalitions, such as heart health or smoking cessation programs; several were of efforts to promote primary care in medical office practices. Others had quite diverse contexts, including one that examined the long-term influences of a 4-day training program for nurses on quality improvement methods and another that assessed projects that had promoted multidisciplinary, community-based education for health professionals. Those that had external funding had received it for periods ranging from 18 months to 8 years; in five sets of programs, external funding was provided for 3 years or less. Six studies were of projects that received no external funding or did not report its extent, for example, when only training was provided to the intervention sites, or multiple sources were used to fund the same intervention, as noted in Table 3.

It would have been desirable to examine systematically the implications for the extent of sustainability of a number of characteristics of the programs reviewed, such as their source(s) of initial funding, differences in program content areas and intervention strategies, and variability in organizational characteristics (e.g., whether the organization itself was a small, struggling nonprofit versus a larger, established organization). Unfortunately, these studies often did not include these detailed descriptors about the sites they assessed, particularly not in terms that were consistent enough across studies to permit comparisons. Furthermore, the original authors' operational definitions of what was meant by sustainability were often quite generous (such as "Are any activities from the project still remaining?") or even nonexistent.

A positive finding from this review is that a substantial proportion of studies found that some type of sustainability was achieved within a majority of the sites studied, as shown in Table 3. I adopted an inclusive definition of sustainability for this tabulation: If the original author stated that the project and/or some of its activities still existed, I coded it as sustained, using studies as the unit of analysis for the frequencies reported below.

- Fourteen of 17 relevant studies reported that 60% or more of the sites showed some sustainability, for at least some activities or the continued existence of community coalitions. (I choose the 60% benchmark as a level showing success in sustainability for at least a majority of the sites studied, but not so high a criterion as to be unattainable. Specific percentages of sites sustained are shown in Table 3.[7])
- Two other studies selected the sites to be observed to show a range of sustainability (Glaser, 1981; Goodman & Steckler, 1989), so the overall percentage sustained was not relevant (both used case study methods).
- Only three studies reported less than 60% sustainability for all of the components studied, all lower than 40% of their sites sustaining. One of these examined the continued use of a short training program 4 years after delivery and found a quite respectable 39% of nurses still using the training.

Many of these studies used exploratory methods to see if anything remained several years after external funding had ended. The fact that so many of them found evidence that project components or activities still existed is suggestive evidence that some form of sustainability is often possible, although certainly not guaranteed. Nevertheless, much greater rigor is needed in future studies concerning the definition and measurement of extent of sustainability. As a minimum, the components or activities in the original project should be detailed, including questions about the specific components that survived or were abandoned and why.

Comparisons Among Types of Measures

Several studies examined more than one type of sustainability measure and found differential results. For example, Shediac-Rizkallah, Scheirer, and Cassady (1997) found that 64% of 28 hospitals reported maintaining some components of their breast cancer screening programs after funding ended, but they provided only 44% as many screening mammograms as during the same period the prior year. In this case, the outcome of benefits sustained for clients was not nearly as high as the percentage of hospitals continuing some activities. Several follow-up studies were conducted of the federal Community Intervention Trial for Smoking Cessation (COMMIT) community trial to promote smoking cessation. One year after the funding ended for the intervention activities, Lichtenstein, Thompson, Nettekoven, and Corbett (1996) found that 9 of 11 (82%) of the intervention communities still had organized coalitions, but they averaged only 68% of their earlier scope of activities. Two years after funding ending for the same COMMIT trial, Thompson, Lichtenstein, Corbett, Nettekoven, and Feng (2000) found that 9 of 11 communities in both the intervention and comparison arms of the original trial maintained coalitions or other organized tobacco control structures, and the activity levels in five areas of intervention were similarly moderate to high in both sets of sites. This study suggests that the continued activity levels for some types of programs might be due to the underlying trends for promotion of such activity in all communities, rather than differential maintenance just in the intervention sites.

Most of these studies did not differentiate among the three types of sustainability measures described by Shediac-Rizakallah and Bone (1998) in the framework described above: the sustainability of beneficial outcomes for clients, the continuation of program activities, and the maintenance of community attention to the problem addressed by the program. As indicated in the final three columns of Table 3, only two studies measured benefits to clients, although three others included some examples of the numbers of clients that continued to be served. One rigorous study (Stange, Goodwin, Zyzanzki, & Dietrich, 2003) followed up a program of interventions to increase the use of prevention services in 37 family medical practices. Using data extracted from patient medical records at least 12 months after the interventions, the researchers found no statistically significant reductions in the overall rates of services to clients, which had increased significantly during the intervention. In other words, services were maintained on average, at the same rates for at least a year following the interventions. Unfortunately, this study did not report or analyze the extent of variability among the 37 medical practices in the study, in the rates of preventive services provided. This follow-up study also did not include data from the original comparison group of medical practices, which were offered a delayed version of the interventions.

Nearly all studies (18 of the 19 studies reviewed) examined whether program activities were sustained, although for some studies, this meant only a single question in a questionnaire or interview asking the respondent whether the program was maintained. Several other studies considered sustainability from an institutionalization perspective and measured a number of

indicators of whether full institutionalization had taken place. Further analysis of influences on the level of sustained program activities is provided below.

Six studies explicitly examined the maintenance of community coalitions or other community structures for continuing to address their problem area (Bracht et al., 1994; Herrera, 2002; Lichtenstein et al., 1996; Lodl & Stevens, 2002; Paine-Andrews, Fisher, Campuzano, Fawcett, & Berkley-Patton, 2000; Thompson et al., 2000). In five of these six studies, at least 60% of their community coalitions continued to operate when the follow-up study was conducted (see Table 3 for details). Few of the other studies assessed whether the program had enhanced or sustained influences on community capacity. As indicated in Table 3, seven additional studies provided some commentary or analysis about the importance of community support for the program activities undertaken but did not attempt to measure or analyze community capacity.

Several studies that examined the sustainability of activities promoted by community coalitions did report on the continued existence of the coalitions as coordinating structures. For example, the Minnesota Heart Health Program (Bracht et al., 1994) used community boards in three cities; two of the three cities incorporated the boards following the program to assume responsibility for maintaining the programs. The community programs analyzed by Paine-Andrews et al. (2000) also developed local advisory boards or steering committees, but the authors did not report how much of this structure remained following the program funding. Two studies of the national COMMIT trial on smoking cessation both report that 9 of 11 intervention communities still maintained tobacco control structures 1 year (Lichtenstein et al., 1996) and 2 years after (Thompson et al., 2000) the end of federal funding. Thompson et al. (2000) further analyzed the "strength" of these coalitions in terms of their independence of agendas, their receipt of funding, and the extent of paid staff members. However, these studies of coalitions do not examine whether or how the continued operation of the coalition structure contributed to increased community capacity, for example, to develop or operate other programs.

One study (Jackson et al., 1994) that did explicitly examine the capacity of communities to carry on health promotion activities was an extended follow-on activity to the Stanford Five-City Project, an interventional research program for heart health led by researchers from Stanford University in the 1980s (Farquhar et al., 1985, 1990). During the 6-year intervention period, the researchers and community participants planned for maintaining the comprehensive program of heart health promotion activities after federal funding ended. The plan involved a networking strategy under a new umbrella nonprofit agency, overseen by a community advisory board, which would locate and disseminate program ideas to local participating agencies. After 3 years of attempting to work within this structure, the participants dissolved it, finding that it resulted in increased conflict among agencies and increased competition for resources and staff time. Subsequently, they shifted to a capacity-building strategy for local health education staff members partnering with the university staff members to develop agency members' skills in program management, grant writing, project evaluation, and other topics requested by the health agency participants. This activity was maintained for 2 years and was viewed as successful by the participants and authors of this descriptive case study. This report ends by defining sustainability as capacity building of individual staff members within their agencies rather than measuring the extent to which they continuing specific activities or achieved outcomes started under the original program. (For this reason, this report was not coded in the tables for this review: It did not use any of the three definitions to measure sustainability. Future research on sustainability and related topics might well include more attention to capacity building as a possible outcome of short-term programs.)

Factors Related to Extent of Sustainability

Most studies in this review provided some explanation for the levels of sustainability they found. They asked, What factors help programs to sustain themselves after external funding ends? Are there some identifiable factors that could help increase the extent of sustainability in the future? This review classifies these factors according to the framework of potentially influential factors suggested by Shediac-Rizkallah and Bone (1998), as shown in Table 4. However, the authors of the studies reviewed did not usually use this framework when suggesting influences on sustainability. Instead, they often took an inductive approach of describing differences between high and low or nonsustained sites, or they obtained participants' perceptions of helpful or detrimental factors. Whenever a study mentioned the positive influence of a factor included in Table 4, it was coded as "yes." (These were sometimes statistically derived associations, such as correlations, but not always. As indicated in Table 2, only eight of the studies used any type of statistical tests.) The many blank cells in Table 4 indicate that the studies made no mention of that factor. In only a few studies was one of these factors examined but found to be not related to the extent of sustainability (coded as "NR" for "not related"). Coding these factors thus required frequent judgment on whether the factors being discussed in a particular study are the same ones mentioned in the framework. For example, is a "strong executive director" (found to be an important factor by Herrera, 2002) the same as having a program champion? I thought these similar enough to be included as "yes" in that column.

Seventeen of the 19 studies provided some analysis or discussion of factors believed to influence sustainability, as shown in Table 4. The unit of analysis for this summary is thus the studies reviewed, not the sites within studies. Few studies used a multivariate method (such as multiple regression) to examine the associations between hypothesized influences and whether the project was sustained. Furthermore, few studies provided operational definitions of their predictor variables or provided information about the variability among sites for those variables. Therefore, cross-study examination of influential factors using the methods of statistical meta-analysis was not possible in this review. Many of the studies stated they used exploratory methods, such as fairly open-ended telephone interviews with project directors, and made little attempt to build on the conceptual frameworks provided by prior research. Therefore, the findings from this review should be viewed as suggestive rather than conclusive.

Influences on sustainability found or mentioned most often were the following:

Concerning project design and its characteristics:

- Twelve studies suggested that programs that were modifiable at the local level were more likely to be sustained. They reported that local organizations were likely to make changes in the initial program design to address a greater perceived need (such as preventing teen smoking rather than adult smoking cessation) or to make it easier to deliver locally. Few of these studies examined whether these modifications were made in essential components of the original program (i.e., was the essence of the original program sustained?). Furthermore, one study (Stevens & Peikes, 2004) noted that frequent changes to meet the priorities of new funders could lead to a loss of focus on the initial program goals.
- Five studies explicitly mentioned the use of volunteers or other low-cost ways of delivering services as a key strategy for sustainability.
- Only four studies mentioned the use of evaluation data as an important vehicle for gaining support needed for continuation. In one of these (Stange et al., 2003), data feedback was an integral component of the intervention strategy promoting the use of prevention services in medical care. Several other programs were focused on interventions whose efficacy was already well established, such as breast cancer screening or smoking cessation. A recent article (Stevens & Peikes, 2004) found that although nearly all of the projects studied had evaluations, only about half of the project directors

interviewed thought that evaluation results contributed strongly to sustainability. Another study (Lichtenstein et al., 1996) noted that most of the intervention communities had already decided to continue before the results of the national evaluation were available. However, program staff members' perceptions of program benefits were found by many studies to contribute to sustainability (see below), even if these benefits were not confirmed by research or evaluation.

Concerning aspects of the organizational setting:

- Thirteen studies emphasized the important role of a program champion, sometimes the executive director. This is a person who is strategically located to have access to upper management as well as influence on, or control over, day-to-day program operations. The champion often enthusiastically advocated for the needs of the program, particularly to help secure resources for its continuation.
- Only four studies cited the strength of existing organizational capacity as a key aspect influencing sustainability. Another strongly empirical study (O'Loughlin et al., 1998) included this variable in its regression analysis and did not find it to be an important influence on the sustainability of 189 heart health interventions at 30 sites. Furthermore, another study (Harris, Henry, Bland, Starnaman, & Voytek, 2003) even found that one aspect of the existing organizational structure, the "silo" structure of academic organizations, was a negative influence on a program to infuse a multidisciplinary and community orientation into medical and nursing education.
- Twelve studies emphasized the "fit" of the new program within the existing organizational mission and/or its standard operating procedures as a key influence on sustainability. Project activities that could be "sold" as contributing to the organization's goals were more likely to receive internal support and even resources that allowed them to be sustained. Furthermore, project activities that could readily fit into existing tasks and procedures were more likely to have the support of operating staff members. However, it is also possible that some of these were continuations of activities that the organization had started up before the "new" funding for the project studied and would have continued even in the absence of that specific source of funding.
- Similarly, 12 studies noted that when staff members or key stakeholders could perceive benefits to themselves and/or to clients, the program was more likely to be sustained. However, such benefits were not necessarily documented by formal evaluation or prior research.

Concerning aspects of the community environment surrounding the program:

- Twelve studies emphasized the key roles played by support from other organizations in the environment, for example, for in-kind resources such as expert advice in fund-raising, for political support, or to help mobilize clients to advocate for new funding.
- Nine studies emphasized the role of funding from other sources, particularly the availability of a larger number of funding sources or the transfer of support to local governmental sources. However, two studies (O'Loughlin et al., 1998; Scheirer, 1990) both found that the actual availability of new funding was not a predictor of sustainability (both used relatively rigorous multiple regression analysis) but that funding was perceived by respondents to be very important to continuation. Some studies appeared to assume that securing additional funding was nearly synonymous with sustainability and did not consider it as a separate factor.

Keeping in mind the methodological limitations of these studies noted above, there is reasonable convergence here on the importance of five factors: (a) The program itself is modifiable over time, (b) the key roles of a program champion, (c) a substantial fit with the underlying organization's mission and procedures, (d) benefits to staff members and/or clients that are readily perceived (but not necessarily documented via formal evaluation), and (e) the importance of support from other stakeholders in the community. Although alternative funding was not explicitly cited as a factor by as many studies, in many reports, new funding was assumed to be needed for sustainability, and other influences were linked to their roles in helping the program to secure new funding. Only a few of these studies were influenced by the prior work of Goodman and Steckler (1989) or Shediac-Rizkallah and Bone (1998) to develop hypotheses about

which variables were likely to influence sustainability. Nevertheless, the influences recorded as positive in Table 4 were those supported by the evidence independently collected for each study.

In addition to the factors suggested by the framework of Shediac-Rizkallah and Bone (1998), I attempted to look at the extent of time that had elapsed after funding until the data were collected about sustainability. Does sustainability deteriorate over time? This question could not be fully addressed in this review, because the time when initial funding ended was ambiguous in many of the studies. Others noted that there were varying time periods since the end of funding among the agencies they studied but did not relate this variability in elapsed time to the extent of sustainability. However, three studies found a positive relationship between time since project start-up and sustainability, whereas two others examined time and found no relationship. A positive relationship means that projects that received funding early in the initiative being studied (perhaps "early adopters") were more likely to have sustained their projects than those funded later. These early adopters might have been organizations with more interest in the program, stronger champions, stronger fit of the program with the underlying mission, or higher status on other underlying influences on sustainability than were projects funded later. There may also be some response bias in this finding: Sites that had earlier funding among those in a study but did not continue that project after funding ended may be less contactable for a survey than are sites that sustained the program. This is especially likely to occur if there is turnover of the initial program director or champion, and the implementing organization does not continue the training and support after the initial implementers leave. (The studies reporting the positive relationship with time since start-up did not break out their response rates by time since startup to enable examining this potential nonresponse bias.)

A few additional influences were noted by one or more studies. External technical assistance from program developers or funders was viewed as helpful by three studies. Two studies supported the use of early planning for sustainability. At least one cited each of the following as helpful to the sustainability processes: continuous staff discussion about how to implement and sustain the program, having a paid coordinator to staff the program, and multiple strategies used for obtaining funds. Other negative influences mentioned were staff turnover, medical practice ownership changes, and a low level of implementation early in the project.

Discussion and Implications for Several Audiences

The studies reviewed found a consistent pattern of evidence that at least some type of sustainability is frequently possible, although their methodological limitations were numerous. However, most studies reported that continuation is not guaranteed: Informants providing the data for these studies frequently cited the challenges they had faced in trying to sustain their projects, particularly in securing funding or other resources needed for continuation. Furthermore, the studies' discussions make clear that sustainability is a continuously evolving process in the life cycle of a project, which begins before the end of initial funding and is not always ensured, even with resources that extend for 1 or 2 years after the end of initial funding.

In many of these studies, sustainability was found to be influenced by a coherent set of factors primarily related to its organizational context and the people behind it, both within and outside the implementing agency. Organizational factors are shown in the importance of the fit with an embedding organization's mission and procedures, as well as the extent to which the program can be modified to adapt to the organization. The importance of leadership and staffing was shown by the fact that more than three fourths of the studies that examined influences on sustainability cited the importance of a champion, someone who is strategically placed within an organization to advocate effectively for the program. A belief in the benefits provided by the program by both staff members and external stakeholders was cited more often than a positive

influence from actual evaluation findings. Equal numbers of studies found that other organizations and community supporters played a key role in helping secure resources and mobilizing support for continuation.

These studies examined a wide range of program types, including community coalitions, community-based prevention programs, training programs, interventions within medical practices, and the use of volunteers to provide community services. The extent of sustainability does not appear to be related to the type of program implemented among these diverse health-related studies. Nevertheless, the question of whether sustainability is easier for some types of programs certainly deserves further attention in future research. Because this review was limited to empirical studies that had focused on health-related programs, the generalizability of its findings to other content areas is unknown. It would be desirable to conduct similar reviews of sustainability studies in different content areas that use short-term funding for demonstration or "seed money" projects, such as education, criminal justice, or social services. One could hypothesize that the extent of sustainability in other sectors is likely to be related to the presence or absence of the organizational and individual supports for it and/or the extent to which the implementing organizations have characteristics similar to the organizations in the health studies rather than differing by the content area of the programs. This hypothesis remains to be tested.

Long-term outcomes other than project sustainability within the funded site might also occur from demonstration programs; other potential outcomes deserve more evaluative attention. For example, this review did not include evaluations of efforts toward long-term change in the ongoing activities within health delivery organizations, often termed quality improvement projects. Other types of social change that might show positive outcomes from demonstration programs include increasing the capacity of the funded organization to implement other new projects, transferring or disseminating the program activities to other organizations, or enhancing the capacity of an entire community to address its social needs. Few of the studies reviewed here addressed any of these alternative outcomes of the programs they were studying.

Operationalizing Sustainability

This discussion of the broad array of potential outcomes for the demonstration projects that were the focus of these studies raises fundamental issues concerning the definition and measurement of sustainability. This review focused on three definitional measures of sustainability: (a) continuing to deliver beneficial services (outcomes) to clients (an individual level of analysis); (b) maintaining the program and/or its activities in an identifiable form, even if modified (an organizational level of analysis); and (c) maintaining the capacity of a community to deliver program activities after an initial program created a community coalition or similar structure (community level of analysis). Other definitions or components of sustainability are also possible, which might focus on other long-term outcomes listed above. Only a few of the studies reviewed here discussed in any detail their own operational definitions of sustainability, a notable exception being the careful work by Goodman and his colleagues on the institutionalization of a program (Goodman & Steckler, 1989; Goodman et al., 1993; Steckler & Goodman, 1989). Most other studies did not address more detailed measurement issues about what level of success or what proportion of a project's original activities, at what level of intensity, need to be present before considering a site as "sustained." Furthermore, if the adaptation of components is viewed as desirable at the local level, at what point is it no longer the "same" program? Clearly, the findings of this review might have been very different if I had restricted the review to any one of these definitions for measuring sustainability or had included only articles with strong methodologies.

Ideally, a logic model or careful process evaluation would be available to define what program activity components are essential to achieve a given outcome. In this case, the maintenance of these components would constitute a good operational definition of program-level sustainability. However, in many instances, the program components have not been carefully defined or even fully implemented before sustainability is assessed, so this guidance on operationalizing sustainability would be difficult to implement. Furthermore, this definition would complicate multiple-site evaluation research about the sustainability of projects with different local components, (e.g., those studied by Stevens & Peikes, 2004, and O'Loughlin et al., 1998). Such inquiries would need to first identify the program components in each site rather than asking only generic questions across many sites about whether each project had continued. For these reasons, documenting the continued extent of beneficial services or outcomes for clients may be the most rigorous definition of sustainability, but it too may be unfeasible unless the local project has collected and continues to collect this type of outcome data.

Many of these issues need further careful evaluative research to sort out. As indicated in the Findings section about the methodologies used by these researchers, future research needs to be explicit about operational definitions of concepts being measured and to report fully the methods used in each study for both the outcome variable, sustainability, and measures of potential influencing factors. Evaluators should build on the work done previously about sustainability, such as the conceptual framework of Shediac-Rizkallah and Bone (1998), Yin's (1979, 1981) routinization dimensions, the extensions of Yin's work by Goodman and others, as well as this review. Evaluation researchers conducting studies of sustainability need to have some background in the literature on organizational behavior to understand the organizational influences that operate across the life cycle of project start-up, implementation, then sustainability and other potential longer term outcomes.

Funding Research on Sustainability

Many issues concerning the scope and rigor of future evaluations of project sustainability depend on the adequacy of financial support for these studies. At the least, funders should continue to provide support for evaluation to go beyond the usual focus on ascertaining effectiveness to grapple with these longer term issues of sustainability. Several studies reviewed here stated they were "exploratory" or had limited funding, primarily to find out whether anything was maintained of programs funded some years earlier and, if so, what factors seemed to influence or improve sustainability. In essence, many studies aimed to advise program funders on "how to do it better," rather than to contribute to a research-based understanding of sustainability. Immediate needs of funding agencies for some evidence about how to foster sustainability may conflict with the likelihood that the research will include rigorous methodology, which usually costs more. However, given the findings reported here that the sustainability of such projects often is possible and that the prior studies suggest a set of factors that are likely to influence these processes, it would be desirable for future sustainability studies to build on and go beyond the results reported here rather than to repeat the same type of exploratory studies.

Some methodological caveats and limitations of this review are worth noting. First, most of the studies used mail or telephone surveys to collect data. Although the respondents to these surveys were likely to report accurately on the existence of a continued project, they may have some limitations in their reporting on the continuation of all its components and activities, particularly if no list of initial project components was available to the researchers. The factors survey respondents reported as influential on sustainability may reflect their assumptions and "hunches" about contributing factors, rather than the findings from independently measured

variables associated with a measured extent of sustainability. Second, few of these studies used on-site observations to measure sustainability. One set of case studies with a careful definition of institutionalization (Goodman & Steckler, 1989; Steckler & Goodman, 1989) rated only 1 of its 10 sites as having a "high" level of institutionalization, with 3 others rated "moderate" and the remaining 6 as "low" institutionalization. This was a lower level of sustainability than reported by most other studies with less careful methodology; however, these sites were selected for case studies from sites expected to show a range of institutionalization progress.[8] Third, only 2 of these studies measured the continuation of benefits for new clients: One (Shediac-Rizkallah et al., 1997) found a greatly reduced level of mammograms during an unfunded period, and the other (Stange et al., 2003) reported no significant reduction in prevention services provided by medical practices by 12 months after the intervention. Thus, the continuation of the same level of services in relation to client benefits cannot be assumed, even when a high level of program activity is reported. Finally, this review of factors associated with sustainability often used judgment in coding an article's narrative discussion into the major categories shown in Table 4. Another reviewer might have started with different categories or coded the source materials differently.

Recommendations

This review has some important implications for several groups of stakeholders who are interested in enhanced sustainability: developers at the local level, external funders of these programs, and evaluators.

Developers at the local level. For local program developers (and evaluators working in a program planning stage), several recommendations can be offered to increase the likelihood of program continuation:

- Choose programs and interventions that relate strongly to your agency's mission and culture, so that support from upper management will be likely, and tasks needed to implement the program will fit within the workloads of available staff members.
- If the program components have been developed elsewhere, engage in thoughtful modifications of components to fit the new organizational context, without destroying the core components contributing to the effectiveness of the original design.
- Identify and support a program champion to take a leadership role in both initial program development and planning for sustainability.
- When designing and publicizing the program, emphasize its benefits for various groups of stakeholders, including staff members and clients, as well as its fit with the major objectives of potential external funders.
- Consider the possible advantages of "routinizing" the program into the core operations of an existing agency rather than continuing it as a "stand-alone" program. Use Yin's (1979, 1981) list of factors contributing to routinization (see Table 1) as a checklist of organizational aspects to work on.

External funding agencies. Funding agencies and their evaluators with interests in sustainability (whether a foundation, a governmental agency, or another initial funding source) may be able to influence this long-term outcome by doing the following:

- Funding projects in existing agencies with some capacity to support them and to provide the expertise needed for carrying out the many facets of sustainability. Or if a new project involves creating a new organization, allow time and resources for building the capacity of that new entity.

- Funding smaller scale projects that also have some local resources involved in them, to build ownership of the project among local stakeholders.
- Identifying, working with, and strengthening local champions to provide the leadership and knowledge of local organizations needed to keep a project going over time.
- Recognizing that programs do not remain static at the local level but are adapted to fit local priorities and capacities. The widely used model of "develop, validate by evaluation, then disseminate" to yield effective programming may not be useful for improving practice if new users modify the program components substantially.
- Allowing enough time and resources in the initial project for it to fully develop its capacity and fully implement the intended program activities.
- Encouraging planning for sustainability early in a project's life cycle, particularly if the project is not the initial research testing the effectiveness of a new program idea. If the project is developing and testing a new program idea, it should include rigorous evaluation and enough time after the results of the evaluation are known to plan for the sustainability of program interventions shown to be effective.
- Evaluation funders should support studies of sustainability even after the initial program funding is terminated, including data collected to assess the continuation of benefits for intended clients. Simply inquiring whether the program continues to exist does not address whether it continues to provide the same scope or types of activities or the same extent of benefits for clients.

Evaluators and researchers. Finally, the methods and findings of this article suggest some implications for evaluators and other researchers investigating sustainability:

- The topic of sustainability requires its own evaluation, apart from and usually after, an evaluation has shown positive results for the program intervention itself. This is part of the life cycle of program evaluation associated with the different stages of program development and delivery.
- Researchers publishing articles about sustainability should be sure to fully document their methods for data collection and analysis, so that the likely validity of their findings can be assessed in relation to the methods used in each study.
- Methods for studying sustainability call for further development of standard ways of operationalizing sustainability and the factors influencing it, so that results of studies can be compared and accumulated by review articles such as this one.
- Studies of sustainability should make greater use of methods to reduce potential bias in findings, such as contacting multiple respondents to obtain convergence in reports of organizational processes and using multiple sources of evidence to triangulate findings.
- Future studies of potential influences on sustainability should start from a broad conceptualization of these factors, for example, drawing explicitly from the frameworks of Shediac-Rizkallah and Bone (1998), Yin (1979, 1981) and Goodman and his colleagues. Limiting future evaluative research to the five influential factors found in this review would be premature, because few of these studies started with a fully operationalized conceptual model of potential factors; few tested these associations statistically.
- The timing of evaluation findings is often too late in the project life cycle to be useful in promoting sustainability; evaluation could be more useful if it included continuously accumulated data about major outcomes, so that interim data about outcomes would be available before the initial funding ends.

This article has examined the available strands of empirical literature about the sustainability of health programs to summarize what has been learned to date about this complex topic. Using a life cycle perspective grounds the study of sustainability in the context of the prior processes surrounding the program's development and implementation. Drawing on the framework suggested by Shediac-Rizkallah and Bone (1998), I found that only a few studies had measured the extent to which beneficial services or outcomes for clients were sustained. Most of the studies

focused only on the sustainability of program activities and presented a reasonably positive picture: 14 of the 17 studies (for which the percentage of sustainability was presented or could be calculated) found that 60% or more of their sites sustained at least some of their activities. Studies that had examined community coalitions also reported that most coalitions continued to exist after external funded ended. However, these findings might have been quite different if the studies reviewed had used a different definition and measurement of the key outcome variable of sustainability or had used more rigorous methods of data collection and analysis. The studies also showed substantial convergence on major influences that were related to increased program maintenance, including programs with modifiable components, an active program champion, a good fit with the implementing organization's mission and procedures, benefits from the program that are visible to staff members and other participants, and support from other community agencies and members. Many of these studies also suggested that these influencing factors are important factors in generating continued financial support. In sum, this body of evidence supports the tentative conclusion that local project sustainability is possible, under the right conditions generated by the convergence of internal and external supporting factors.

Notes

1. See further discussion of the units of analysis and data definitions used in these studies in the section titled "Conceptual Frameworks."

2. Both of these frameworks were published after the review and coding were completed for this article, but the variables they included in their conceptualizations do not differ in major ways from Shediac-Rizkallah and Bone's (1998) framework.

3. See further discussion and examples in the section titled "Extent and Types of Sustainability."

4. One of these studies (Wallin, Bostrom, Wikblad, & Ewald, 2003) appropriately contacted only one person per intervention, because the intervention was a training program for individuals whose long-term sustained use was being assessed.

5. See the section titled "Extent and Types of Sustainability" on funding sources. Some interventions were studied that did not have a single source of external funding (e.g., the school-based Fluoride Mouth Rinse Program described in Scheirer, 1990), or multiple interventions were studied that had funding from diverse sources (Evashwick & Ory, 2003).

6. For the most part, these studies did not report how much variability there was across sites for the predictor variables assessed. As pointed out by an astute reviewer, a factor that was relatively homogeneous across the sites in a specific study would not be supported as a predictor variable, or influence, on the sustainability outcome.

7. Only 5 studies reported 80% or more of their sites sustained; see Table 3 for specific percentages for each study or to calculate the extent of sustainability at different cutoff points.

8. This study was not included in the tabulation of overall sustainability achieved, because its sites were selected on the basis of this outcome variable.

References

Altman, D. (1995). Sustaining interventions in community systems: On the relationship between researchers and communities. *Health Psychology*, *14*, 526-536.

Barab, S. A., Redman, B. K., & Froman, R. D. (1998). Measurement characteristics of the Levels of Institutionalization scales: Examining reliability and validity. *Journal of Nursing Measurement*, *6*, 19-33.

Bracht, N., Finnegan, J. R., Jr., Rissel, C., Weisbrod, R., Gleason, J., Corbett, J., et al. (1994). Community ownership and program continuation following a health demonstration project. *Health Education Research*, *9*, 243-255.

Elder, J. P., Campbell, N. R., Candelaria, J. I., Talavera, G. A., Mayer, J. A., Moreno. C., et al. (1998). Project Salsa: Development and institutionalization of a nutritional health promotion project in a Latino community. *American Journal of Health Promotion*, *12*, 391-401.

Evashwick, C., & Ory, M. (2003). Organizational characteristics of successful innovative health care programs sustained over time. *Family and Community Health*, *26*, 177-193.

Farquhar, J. W., Fortmann, S. P., Flora, J. A., Taylor, C. B., Haskell, W. L., Williams, P. T., et al. (1990). Effects of community-wide education on cardiovascular disease risk factors: The Stanford Five-City Project. *Journal of the American Medical Association*, *264*, 359-365.

Farquhar, J. W., Fortmann, S. P., Maccoby, N., Haskell, W. L., Williams, P. T., Flora, J. A., et al. (1985). The Stanford Five-City Project: Design and methods. *American Journal of Epidemiology*, *122*, 323-334.

Glaser, E. M. (1981). Durability of innovations in human service organizations. *Knowledge: Creation, Diffusion, Utilization*, *3*, 167-185.

Goodman, R. M., McLeroy, K. R., Steckler, A. B., & Hoyle, R. H. (1993). Development of Level of Institutionalization (LoIn) scales for health promotion programs. *Health Education Quarterly*, *20*, 161-178.

Goodman, R. M., & Steckler, A. (1989). A model for the institutionalization of health promotion programs. *Family and Community Health*, *11*, 63-78.

Goodson, P., Smith, M. M., Evans, A., Meyer, B., & Gottlieb, N. H. (2001). Maintaining prevention in practice: Survival of PPIP in primary care settings. *American Journal of Preventive Medicine*, *20*, 184-189.

Green, L. W. (1989) Comment: Is institutionalization the proper goal of grantmaking? *American Journal of Health Promotion*, *3*, 44.

Harris, D. L., Henry, R. C., Bland, C. J., Starnaman, S. M., & Voytek, K. L. (2003). Lessons learned from implementing multidisciplinary health professions educational models in community settings. *Journal of Interprofessional Care*, *17*, 7-20.

Hawe, P., Noort, M., King, L., & Jordens, C. (1997). Multiplying health gains: The critical role of capacity-building within health promotion programs. *Health Policy*, *39*, 29-42.

Herrera, C. (2002). *Report on sustainability in Faith in Action* (Draft report to the Robert Wood Johnson Foundation from Public/Private Ventures).

Hogg, W., Baskerville, N., Nykiforuk, C., & Mallen, D. (2002). Improved preventive care in family practices with outreach facilitation: Understanding success and failure. *Journal of Health Services Research & Policy*, *7*, 195-201.

Jackson, C. J., Fortmann, S. P., Flora, J. A., Melton, R. J., Snider, J. P., & Littlefield, D. (1994) The capacity-building approach to intervention maintenance implemented by the Stanford Five-City Project. *Health Education Research*, *9*, 385-396.

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.

Labonte, R., & Laverack, G. (2001). Capacity building in health promotion, Part I: For whom? And for what purpose? *Critical Public Health*, *11*, 111-127.

Lichtenstein, E., Thompson, B., Nettekoven, L., & Corbett, K., for the COMMIT Research Group (1996). Durability of tobacco control activities in 11 North American communities: Life after the Community Intervention Trial for Smoking Cessation (COMMIT). *Health Education Research*, *11*, 527-534.

Livit, M., & Wandersman, A. (2004). Organizational functioning: Facilitating effective interventions and increasing the odds of programming success. In D. M. Fetterman & A. Wandersman (Eds.), *Empowerment evaluation principles in practice* (pp. 123-154). New York: Guilford.

Lodl, K., & Stevens, G. (2002). Coalition sustainability: Long term successes & lessons learned. *Journal of Extension*, *40*, 1-8.

Mancini, J. A., & Marek, L. I. (2004). Sustaining community-based programs for families: Conceptualization and measurement. *Family Relations*, *53*, 339-347.

Newcomer, K. E., Hatry, H. P., and Wholey, J. S. (1994). Meeting the need for practical evaluation approaches: An introduction. In J. S. Wholey, H. P. Hatry, & K. E. Newcomer (Eds.), *Handbook of practical program evaluation* (pp. 1-10). San Francisco: Jossey-Bass.

O'Loughlin, J., Renaud, L., Richard, L., Sanchez-Gomez, L., & Paradis, G. (1998). Correlates of the sustainability of community-based heart health promotion interventions. *Preventive Medicine*, *27*, 702-712.

Paine-Andrews, A., Fisher, J. L., Campuzano, M. K., Fawcett, S. B., & Berkley-Patton, J. (2000). Promoting sustainability of community health initiatives: An empirical case study. *Health Promotion Practice*, *1*, 249-258.

Pluye, P., Potvin, L., & Denis, J. L. (2004). Making public health programs last: Conceptualizing sustainability. *Evaluation and Program Planning*, *27*, 121-133.

Scheirer, M. A. (1981). *Program implementation: The organizational context*. Beverly Hills, CA: Sage.

Scheirer, M. A. (1987). Program theory and implementation theory: Implications for evaluators. In L. Bickman (Ed.), *New directions for program evaluation, Vol. 33: Using program theory in evaluation* (pp. 40-68). San Francisco: Jossey-Bass.

Scheirer, M. A. (1990). The life cycle of an innovation: Adoption versus discontinuation of the Fluoride Mouth Rinse Program in schools. *Journal of Health and Social Behavior*, *31*, 203-215.

Scheirer, M. A. (1993). Are the Level of Institutionalization scales ready for "prime time"? A commentary on "Development of Level of Institutionalization (LoIn) scales for health promotion programs." *Health Education Quarterly*, *20*, 179-183.

Shediac-Rizkallah, M. C., & Bone, L. R. (1998). Planning for the sustainability of community-based health programs: Conceptual frameworks and future directions for research, practice and policy. *Health Education Research*, *13*, 87-108.

Shediac-Rizkallah, M. C., Scheirer, M. A., & Cassady, C. (1997, May). *Sustainability of the Coordinated Breast Cancer Screening Program* (Final report to the American Cancer Society). Baltimore: Johns Hopkins University School of Hygiene and Public Health.

Stange, K. C., Goodwin, M. A., Zyzanzki, S. J., & Dietrich, A. J. (2003). Sustainability of a practice-individualized preventive service delivery intervention. *American Journal of Preventive Medicine*, *25*, 296-300.

Steckler, A., & Goodman, R. M. (1989). How to institutionalize health promotion programs. *American Journal of Health Promotion*, *3*, 34-44.

Stevens, B., & Peikes, D. (2004). *When the funding stops: Do grantees of the Local Initiative Funding Partners program sustain themselves?* (Report to the Robert Wood Johnson Foundation from Mathematica Policy Research).

Thompson, B., Lichtenstein, E., Corbett, K., Nettekoven, L., & Feng, Z. (2000). Durability of tobacco control efforts in the 22 Community Intervention Trial for Smoking Cessation (COMMIT) communities 2 years after the end of intervention. *Health Education Research*, *15*, 353-266.

Wallin, L., Bostrom, A. M., Wikblad, K., & Ewald, U. (2003). Sustainability in changing clinical practice promotes evidence-based nursing care. *Journal of Advanced Nursing*, *41*, 509-518.

Wandersman, A., Imm, P., Chinman, M., & Kaftarian, S. (2000). Getting to outcomes: A results-based approach to accountability. *Evaluation and Program Planning*, *23*, 389-395.

Weiss, H., Coffman J., & Bohan-Baker, M. (2002) *Evaluation's role in supporting initiative sustainability*. Cambridge, MA: Harvard Family Research Project. Available at http://www.hfrp.org

Yin, R. K. (1979). *Changing urban bureaucracies: How new practices become routinized*. Lexington, MA: Lexington Books.

Yin, R. K. (1981). Life histories of innovations: How new practices become routinized. *Public Administration Review*, *41*, 21-28.