The Virginia Abstinence Education Initiative Evaluation Structure: A Lesson in How to Successfully Overcome the Challenges of Multi-Site Program Evaluation

Allen Lewis
Virginia Commonwealth University
Richmond, VA

Sarah Jane Brubaker
Virginia Commonwealth University
Richmond, VA

Ariane S. Karph
Virginia Commonwealth University
Richmond, VA

Brian Ambrose
Virginia Commonwealth University
Richmond, VA

The Virginia Abstinence Education Initiative Evaluation Structure: A Lesson in How to Successfully Overcome the Challenges of Multi-Site Program Evaluation

Allen Lewis and Sarah Jane Brubaker
Virginia Commonwealth University

Gale Grant and Kathy Wibberly
Virginia Department of Health

Abstract: This discussion provides an overview of the evaluation process of the Virginia Abstinence Education Initiative (VAEI). It details the basic principles that premise the evaluation structure. The evaluation structure utilized by the VAEI is an intentional one, designed to provide the most rigorous approach possible in order to have maximum confidence in the quality of the data produced by this statewide, multi-year effort. The authors argue that this type of informed approach grounded in a high degree of evaluation rigor can help to overcome the challenges typically associated with multi-site program evaluation.

Introduction

The Virginia Abstinence Education Initiative (VAEI) is a statewide, multi-year effort to implement new educational approaches to help youth develop skills necessary to delay sexual involvement, and to evaluate systematically the effectiveness of those approaches. Funding for this effort is federal Title V Block Grant dollars. While the conduct of rigorous evaluation was not a requirement of receiving this funding, Virginia decided to emphasize program evaluation to assist programs in maximizing impact on participants.

From the outset, the Virginia Department of Health (VDH) acknowledged the importance of the VAEI being data-driven as well as structured through a multi-site approach. To successfully achieve both of these goals would require an intentional and systematic process grounded in experience. The VDH had experience in managing other multi-site program evaluation

initiatives, and realized the value of seeking experienced evaluation expertise as a partner in the endeavor. To this end, the Virginia Commonwealth University (VCU) Survey and Evaluation Research Laboratory (SERL), an experienced university based applied social science research organization, was contracted to assist with the VAEI.

Usual Challenges of Multi-Site Program Evaluation
Multi-site evaluations (MSEs) are increasingly widespread due to their methodological advantages (e.g., increased generalizability of findings, maximized sample size, efficiency of time and testing of contextual effects) and their response to political and social demands (Cottingham, 1991; Mowbray & Herman, 1991; Turpin & Sinacore, 1991). MSEs are not without their challenges, however, and these challenges impact decisions regarding program design, budget, staffing and other issues.

The majority of challenges to conducting MSEs fall into two categories: 1) gaps and tensions between researchers/evaluators and local program staff/community members; and 2) organizational and administrative challenges across sites. While many of the challenges related to the former are common to program evaluation in general, they are intensified and exacerbated in MSEs. The latter types of challenges are more specific to coordination among multiple sites.

Program evaluations can experience tensions between the two often competing arenas of researchers/evaluators and local program staff/community members (Telfair & Mulvihill, 2000). These conflicts reflect differences and gaps between the groups in terms of: 1) the value placed on research/evaluation and 2) the level of understanding and knowledge about, and skills for conducting program evaluation.

Program staff and community members often do not place as much value on program evaluation as do researchers for several reasons. First, some program and community members distrust and/or fear evaluation efforts that may emphasize negative aspects of the program. In addition, limited resources often create a competitive atmosphere in which greater priority is placed by program staff on providing services than on evaluation efforts. This conflict is related to financial resources, staff time, or both. Often program staff members feel that the evaluators identify outcome requirements that conflict with or overlook local program goals, objectives or strategies. Finally, program staff and community members may not always be convinced that their programs will benefit from evaluation efforts. Judd, Frankish and Moulton (2001) describe these conflicts as tensions between empowerment/participation/collaboration and evidence-based decision making, as well as accountability, funding sources' and government decision makers' preoccupation with measuring outcomes.

Constantine and Cagampang (1998) describe how conflicts between program and evaluator priorities can lead to "motivational drift," the lack of motivation toward compliance with evaluation among program staff.   This is typically characterized by feelings of "detachment" or lack of ownership by program staff of the evaluation design, "irrelevance" or insufficient applicability of design to local program needs, and "outcome pessimism" or the belief of program staff that positive results will not be found by the evaluation.

Gaps and tensions between evaluators and program staff in terms of the value placed on evaluation are exacerbated in MSEs. While each site may struggle with one or more issues described above, each may perceive its position differently and feel that its situation is unique.

Obtaining cooperation from the various local sites becomes even more difficult than in single-site program evaluation.

In addition to differences in support for evaluation, there are often gaps between program staff and evaluators in terms of their familiarity with and understanding of program evaluation.   Like the issue of support for/acceptance of evaluation, this type of "terminology/knowledge" gap is exacerbated in MSEs where there may be great variation among sites in terms of their familiarity with, and skill levels for conducting evaluation (Sambrano, Springer, & Hermann, 1997). This poses challenges in terms of staffing and training decisions (Turpin & Sinacore, 1991a). For example, evaluators often grapple with the decision to hire new staff to coordinate evaluation efforts versus utilizing existing local program staff.

The second major area of challenges to conducting successful MSEs relates to the diversity among sites and the administration and management of the data and other evaluation components. Just as variation across sites can exacerbate many of the challenges described above, variation among sites in terms of program implementation poses challenges to a standardized research protocol and ultimately data quality (Constantine & Cagampang, 1998; Cottingham, 1991; Mowbray & Herman, 1991; Ponirakis, 2002; Tushnet, 1995). These differences pose challenges to organizational, procedural and statistical issues (Turpin & Sinacore, 1991b). Research designs must include ways of identifying and analyzing contextual effects. Data collection methods must be standardized and clearly defined and communicated to program and evaluation staff.

All of these challenges – tensions between program staff and evaluators, gaps in terminology/expertise, and variation across sites in terms of program implementation and context – pose threats to fidelity to the implementation model and ultimately to data quality.

Recommendations & Strategies to Overcome the Challenges

Researchers identify various strategies for overcoming challenges to successful MSEs, addressing either or both of the general types of challenges described above. Sambrano et al. (1997) suggest that a combination of decentralized and centralized strategies can be useful in overcoming challenges of MSEs, and these types of strategies can be viewed as addressing the two types of challenges. Specifically, whereas decentralized strategies typically focus on acknowledging, addressing, and valuing individual sites' particular needs and strengths, centralized strategies tend to address the challenges across diverse sites.

Decentralized approaches can help to prevent what Constantine and Cogampang (1998) refer to as "motivational drift," discussed above. Such strategies that address tensions and competing perspectives between local program staff and evaluators often focus on obtaining cooperation, building local support, involving the community, and building trust. Many such strategies are referred to as "inclusive evaluation approaches" and described as participatory, collaborative, or empowering, involving program staff as full partners in the evaluation process. Constantine and Cogampang (1998) review three such approaches, 1) utilization-focused evaluation by Patton, 2) continuous process improvement or total quality management by Deming, and 3) Faulkner's participatory planning model. They suggest that taken together, these three approaches yield three common principles:

- The full range of stakeholders must be meaningfully involved in all aspects of the evaluation;
- An evaluation must be flexible and responsive to local conditions and needs; and
- Data must be regularly and meaningfully shared (Constantine & Cogampang, 1998).

Resnicow and Kirby (1997) similarly recommend designing an evaluation that is collaborative rather than hierarchical, increasing communication between program and evaluation staff, including positive behaviors as outcomes, and involving the community in evaluation. Fetterman (2001) in the empowerment approach to conducting program evaluation, makes clear the strength of evaluations that involve stakeholders in all aspects. Koch, Lewis, and McCall (1998) also describe the benefits of involving stakeholders in designing outcome management systems to serve as infrastructure to support routine program evaluation. Other researchers suggest additional strategies for building local support such as national workshops and regular quarterly or biannual meetings of staff and evaluators to build community, share experiences, and provide feedback. Many argue that essential to the process is the provision of regular communication and returning locally generated data to the sites for their use in assessing and improving local programs. Browne, Clubb, Aubrecht, and Jackson (2001) also suggest that recognizing contributions of community members and local staff and expressing gratitude for their efforts is important. Constantine suggests that providing "regular, immediate, public, and constructive data quality feedback to local-site staff" can help avoid problems related to "motivational drift." Additional decentralized strategies include providing modular survey options for data collection and reporting, multiple evaluation design strategy options, and support for site-specific local evaluation components (Sambrano, Springer, & Hermann, 1997).

Some of the more centralized approaches that assist with organizational and administrative challenges across different sites include implementing a mandatory questionnaire for all sites (Sambrano, Springer, & Hermann, 1997) and providing technical support and assistance (via group meetings as well as on-site visits) to sites that includes a written instructional manual for data collection forms and other standardized evaluation procedures. Although the "decentralized" strategies discussed above can help to overcome the challenge of garnering local and community support for evaluation, many also address some of the organizational and administrative challenges. For example, regular communication, meetings and sharing data all facilitate evaluators' oversight of evaluation and program activities and ability to address and solve problems in a timely manner.

Additional strategies that address the challenges of diversity among sites and potential threats to data quality include focusing on process data and specific program theory (Mowbray & Herman, 1991). Fetterman (2001) similarly emphasizes the importance of evaluations that are theory-driven. These aspects are critical in assessing the validity of the intervention as they provide evidence of the degree to which the program was delivered as planned and why the observed effects were achieved.

This discussion of challenges and strategies illustrates the complexity of these issues and suggests that successfully overcoming such challenges is costly. Browne et al. (2001) similarly suggest that many of the strategies suggested by researchers require intensive investments of time and resources, critical components of a successful MSE.

Method

Virginia Abstinence Education Initiative Evaluation Structure
In response to the usual challenges of multi-site program evaluation, the evaluation structure of the VAEI was developed as part of an intentional process to accomplish the objective of maintaining the integrity of the data management process, and ultimately ensuring the quality

of the evaluation data. To accomplish this, several key principles were embraced from the start. Each principle discussed below addresses one or more of the usual challenges of MSEs and represents a guiding parameter for the VAEI evaluation. Considered together, the principles constitute a framework that strengthened the VAEI evaluation, and increased confidence in the data quality.

Discussion

Building Evaluation in on the Front End

Building evaluation in on the front end represented a realization that to successfully manage a multi-site evaluation process requires a deliberate and informed approach. This approach helped proactively address some of the challenges typically associated with multi-site evaluation such as:

- ensuring that evaluation is theory-driven and incorporated at the point of program inception;
- maintaining fidelity of the intervention across sites, optimizing communication;
- promoting a consensus view on the big picture of the evaluation and its operational aspects;
- strengthening buy-in among all evaluation stakeholders;
- anchoring consistent and predictable evaluation expertise at each site; and
- optimizing data quality.

Each of the above-mentioned challenges associated with multi-site program evaluation was learned experientially in over a decade of program evaluation work conducted by the VCU-SERL with the state of Virginia's adolescent sexual health programs within the VDH and the expenditure of over 25 million dollars. Specifically, both the Virginia Teenage Pregnancy Prevention and the VAEI had structured their multi-site program evaluations over the years to address these challenges.

Ensuring Adequate Evaluation Resources were Available on this Project

One of the key lessons learned from previous experiences by VDH and SERL, and a challenge repeatedly acknowledged in the literature, is to successfully implement a multi-site program evaluation requires adequate resources. To this end, in the VAEI, there were seven distinct roles directly related to program evaluation that are handled by designated staff. Each role addresses one or more of the usual MSE challenges. The roles are: (1) local program data manager, (2) local evaluation consultant, (3) VDH evaluation director, (4) VDH evaluation data manager, (5) SERL evaluation director, (6) SERL evaluation data manager, and (7) VAEI Evaluation Consortium.

Employing a Centralized Evaluation Structure

At VAEI inception, centralization was considered to be a bedrock principle upon which the entire evaluation was based. It was reasoned that only through maximum centralization of all evaluation and data aspects could adequate and proper controls be implemented to ensure confidence in the results. Once the evaluation and data management processes were developed, this information was provided to local program sites as the evaluation operational framework. While local program sites were encouraged to continually provide feedback about the efficacy of the evaluation structure and associated data management procedures, all

conceptual design, development and refinement of data management procedures was controlled centrally by VDH and SERL.

Maximizing Buy-in among All Evaluation Stakeholders

In addition to the strong emphasis within the VAEI evaluation structure on centralized oversight, there was clear recognition early of the importance of having broad and sustainable support and buy-in among all evaluation stakeholders. That support needed to be strong from project inception through the end of the initiative, because, as discussed by Constantine and Cagampang (1998), problems with "motivational drift" tend to increase over time. This was one of the primary strategies employed to avoid typical problems with tensions between evaluators and local stakeholders.

Beyond the groups of stakeholders who had formalized evaluation roles, several more key stakeholders were the targets of efforts to achieve buy-in. Those key stakeholders, without formal evaluation roles, fell into two groups: (1) staff persons charged with delivering the VAEI curricula in local programs and (2) key local program gatekeepers and decision-makers who made the programs possible because they authorized access to youth participants, typically school personnel.

Maintaining a Science-based Focus

Several attributes enhanced the science-based nature of this initiative. They are: use of a limited number of theory-driven, standardized abstinence education curricula, some of which had been replicated and evaluated previously; collection of implementation and impact data; and the use of a rigorous evaluation design (longitudinal focus, pre and posttests, and comparison groups). The threefold combination means this evaluation was highly scientific, especially in comparison to typical program evaluations.

Striving toward Continuous Improvement

The final principle the VAEI evaluation structure embraced is that of continuous improvement. This principle resulted in the recognition that ultimately the evaluation structure must demonstrate steady enhancement over time. That is, the evaluation structure of the VAEI had attempted to meet each local program where it was in the first year of the initiative, with the goal of moving forward toward full actualization of a science-based, rigorous evaluation structure before the end of the project. The following five strategies had been instrumental in the manifestation of the continuous improvement principle.

- Annual updates to the evaluation technical assistance manual.
- Local programs have been given assistance in implementing recommendations associated with evaluation barriers.
- Feedback reports on data quality have been made available to local sites.
- Local evaluation consultants have been required to have regular contact with local program staff.
- The VAEI Evaluation Consortium met quarterly to provide oversight.

Conclusions & Recommendations

Lessons Learned

The major lesson learned from the VAEI experience has been an affirmation of the value of investing a substantial level of resources for program evaluation and utilizing a highly structured evaluation approach. Another lesson learned is that there is no such thing as devoting too

much effort to exploring the comparability of comparison sites. It was also learned that having an evaluator for each local program has been invaluable. Finally, it was learned that having formal written agreements with all participating stakeholders stabilizes evaluation participation over time.

Going forward, it would behoove other adolescent sexual health or any youth development program to take heed of these six principles in designing multi-site program evaluations. Youth development program managers will have to be particularly skilled and agile these days to ensure that youth services are lean, given tight resources, and effective, given increasing demands for program accountability and pressure to achieve key customer outcomes. To do this, they will have to engage in data-driven decision making, a process that involves intentionally, systematically and routinely using data to make decisions (Lewis, Armstrong, & Karpf, 2005), and data-driven decision making is also known as program evaluation.

In this day and time, it is commonplace for program evaluations to involve multiple sites. It is also important to realize that the outcome evaluation movement is a current imperative for professionals working in health, human services, and rehabilitation (Lewis, Armstrong, Taylor, & Spain, 2006). Given the current push around planned and systematic outcome evaluation and performance monitoring, increasingly youth development and other health, rehabilitation and human service programs are looking to develop outcome measurement systems as a tool to manage the multitude of program evaluation activities. Only then, can the final principle, "striving toward continuous improvement" become a catalyst that can have the ultimate effect of bringing to fruition positive youth development outcomes.

References

Browne D.C., Clubb, P.A., Aubrecht, A.M. & Jackson, M. (2001). Minority health risk behaviors: An introduction to research on sexually transmitted diseases, violence, pregnancy prevention and substance abuse. *Maternal and Child Health Journal,* 5(4), 215-224.

Constantine, N. & Cagampang, H. (1998). Improving local evaluation utility within multi-site risk behavior prevention program evaluations. Proceedings of the American Psychological Association Annual Conference. San Francisco, CA.

Cottingham, P. (1991). Unexpected lessons: Evaluation of job-training programs for single mothers. In Turpin, R. & Sinacore, J. (Eds.) *Multisite evaluations* (59-70). San Francisco, CA: Jossey-Bass, Inc.

Fetterman, D.M. (2001). Foundations of empowerment evaluation. Thousand Oaks: Sage Publications.

Judd J., Frankish C.J. & Moulton G. (2001). Setting standards in the evaluation of community-based health promotion programmes: A unifying approach. *Health Promotion International* 16(4), 367-380.

Koch, R., Lewis, A. & McCall, D. (1998). A multistakeholder-driven model for developing an outcome management system. Journal of Behavioral Health Services Research, 25(2), 151-162.

Lewis, A.N., Armstrong, A.J. & Karpf, A. (2005). Using data to improve outcomes in rehabilitation practice. *Journal of Rehabilitation Administration*, 29(1), 43-56.

Lewis, A.N., Armstrong, A.J., Taylor, A.H. & Spain. S. (2006). Determining a vocational rehabilitation program's readiness for outcome-focused program evaluation. *Journal of Rehabilitation Administration*, 30(1), 245-257.

Mowbray, C. & Herman, S. (1991). Using multiple sites in mental health evaluations: Focus on program theory and implementation issues. In Turpin, R. & Sinacore, J. (Eds.) *Multisite evaluations* (45-57). San Francisco, CA: Jossey-Bass, Inc.

Resnicow, K. & Kirby, J. (1997). Finding common ground: on the marriage of evaluation and demonstration programs. *Journal of Adolescence* 17, 10-20.

Rog, D. & Ponirakis, A. (2002). Cross-site evaluation of the CMHS/CSAT homeless family initiative: challenges and opportunities. Proceedings of the 130th Annual Meeting of the American Public Health Association. Philadelphia, PA.

Sambrano, S., Springer, J. & Hermann, J. (1997). Informing the next generation of prevention programs: CSAP's cross-site evaluation of the 1994-95 high-risk youth grantees. American Journal of Community Psychology, 25(5), 375-395.

Telfair, J. & Mulvihill, B.A. (2000). Bridging science and practice: The integrated model of community-based evaluation (IMCBE). *Journal of Community Practice*, 7(3), 37-65.

Turpin, R. & Sinacore, J. (1991). Multiple sites in evaluation research: A survey of organizational and methodological issues. In Turpin, R. & Sinacore, J. (Eds.) *Multisite evaluations* (5-18). San Francisco, CA: Jossey-Bass, Inc.

Turpin, R. & Sinacore, J. (Eds.). (1991). *Multisite evaluations*. San Francisco, CA: Jossey-Bass, Inc.

Tushnet, N. (1995). Toward a general approach to multisite program evaluation. Proceedings of the Annual Meeting of the American Educational Research Association. San Francisco, CA.