

Program Evaluation: Lessons From the Field

Contributed by Vivian R. Bergel, PhD, LSW, and Peggy McFarland, PhD, LSW

Engaging in evidence-based research to support the viability of any program is acknowledged by funders to be vitally important to address such issues as accountability, credibility and, of course, sustainability. If program evaluation is, theoretically, seen as important, why do so few organizations engage in it? The difficulty might be associated with the perceived barriers to conducting such research—barriers that might include time, lack of willing personnel, or lack of knowledge of how to proceed. {mosgoogle right}

The purpose of this article is to provide an example of a program evaluation and, subsequently, explain clearly and concisely how program evaluation can be done in-house by existing personnel. Specific procedures will be addressed that can be followed and replicated. The results of the program evaluation can be used to enhance, refine, publicize, or support the request for grants and awards. The benefits are only limited by the imagination of the board, agency director, and/or staff.

Mika (1996) states that, "to complete a basic, descriptive outcome evaluation...one does not need to be proficient in high-level statistical concepts..." (p. 9). However, the evaluator should have a knowledge of basic descriptive statistics (1996). Many undergraduate and graduate programs incorporate, at least, one research course that usually includes a module on statistics, or requires the completion of a statistics course as a prerequisite or co-requisite to enrolling in the research course. This exposure to statistical and research methodology should provide a foundation for the evaluator to begin.

Gibbs (2003), too, stated that data collection need not be elaborate and time consuming. Time is always a factor and "data need to be simple and related directly to what you are trying to accomplish" (p. 238). Unrau, Garbor, and Gainnell (2007) maintain that outcome evaluation is a practical activity (p. 192) and that if an administrator can run an agency, he or she can direct an evaluation process (2003).

The tasks or steps needed to conduct outcome research are universal throughout the literature (Westerfelt & Dietz, 2005; Unrau et al., 2007; Powell, 2006; Gibbs, 2003; Mika, 2006; United Way of America, 1996).

In general they consist of:

Determining the research questions, i.e., What do you need to know, and/or what part of the program are you trying to evaluate?

Reviewing the literature to support or refute the research question (evidence-based research) and investigate relevant techniques that have proven reliability and validity. Research-based literature can be found on many search engines. If the agency cannot access those search engines, public libraries or colleges and/or universities can prove helpful in this endeavor.

Incorporating a pre-test/post-test measure of program effectiveness. The agency can edit an in-house measure to conform to a Likert scale (for example: very frequently, somewhat frequently, occasionally, not at all) for ease of measurement. We would recommend the addition of a standardized scale to be used in a pre-test/post-test format to reinforce the reliability and validity of the in-house measure. These inventories can be purchased from the scale's author or publishing house.

Evaluating the results through the use of descriptive measures that report on the pre-test and post-test means.

Provide a summary and conclusion noting limitations, if appropriate.

If the agency wishes to examine anecdotal information regarding the program, open-ended questions can be designed to analyze specific content and themes important to the agency. This qualitative research can both reinforce what aspects of the program are successful and what may need to be modified for future participants.

An Example

In 2006, we conducted a program evaluation of an eight-week educational peer support program in Pennsylvania. Our first task was to meet with both the program director and the executive director to determine the agency's needs and objections.

We used both quantitative and qualitative methods to assess the effectiveness of a children's bereavement program. Quantitatively, a pre-post test design was used. We incorporated a Likert scale into the program's "Common Grief Reactions" check sheet. The pretest was part of an intake form that included a parent's report of common grief reactions experienced by his or her child and other changes the child was experiencing. In an effort to maximize the validity of the responses, the Parent Version of the Child Depression Inventory (CDI) (Kovacs, 2003) was also distributed by the program director and completed by the caregivers prior to the onset of the program. The child also completed the children's version of the Child Depression Inventory before beginning the bereavement program. We chose the CDI based on previous research, which attested to the CDI's reliability and validity in assessing the presence of depression in children from ages six to 18 (Kovacs, 2003). We were interested in discovering whether or not any statistically significant or clinical change in depression levels took place following the children's involvement in the program. We saw the CDI as a useful tool to confirm the reliability and validity of the in-house "Common Grief Reactions" instrument.

All of the participants enrolled in a children's bereavement program during an eight-week session completed the pre-test forms on the first night of the program. A parent or guardian completed an informed consent form to participate in

this study. The post-test was administered one month after the completion of the bereavement group. The program director sent the caregivers another copy of the "Common Grief Reactions" check sheet and the parent's and child's Child Depression Inventory forms. The caregivers were asked to complete it and return it in a provided agency addressed and postage supplied envelope.

To conduct the qualitative component of the assessment, we met with nine out of 14 randomly chosen families during a two-month period. These scheduled interviews took place at the family residences in various towns and villages in central Pennsylvania. Questions that we developed were asked of nine caregivers and 13 children. The children's ages ranged from eight to 16. These interviews were conducted in separate areas of the respondents' homes to insure openness and confidentiality. Answers that were both written and recorded audibly were gathered in an attempt to gather anecdotal information regarding specific aspects of the bereavement program. We analyzed these responses for content and themes.

The pre- and post-test comparisons on the grief reaction and depression inventory provided insight regarding the success of the children's bereavement program. Both the grief reaction scale and depression inventory produced positive results when comparing the pre- and post- responses.

Interviews with the children (qualitative responses) provided insight regarding program assessment. For example, one child thought that specific themes were needed to generate focused discussions. The children who were interviewed also felt that there was too much emphasis on discussing negative feelings, and positive feelings needed to be discussed, as well. Discussions including the theme of resiliency may build and/or reinforce the strength and coping ability of the children. Several children mentioned that they would have liked to have had the opportunity to share their thoughts privately with the group leader. They were not always comfortable sharing in the group. The group leader should continuously monitor the quality of the children's interactions to assess whether or not the group experience seems appropriate for all participants. Individual therapy may serve as an important adjunct to the group experience.

Conclusion

It is our hope that the idea of conducting a program level evaluation is seen as important and that the process seems achievable. Evaluating for effectiveness serves to improve service delivery. This, of course, is the goal of every agency.

References

- Gibbs, E. (2003). Evidence-based practice for the helping professions. Pacific Grove, CA: Brooks/Cole.
- Kovacs, M., & MHS Staff. (2003). Children's depression inventory (CDI): Technical manual update. North Tonawanda, NY: Multi-Health Systems.
- Mika, L. (1996). Program outcome evaluation: A step-by-step handbook. Milwaukee, WI: Families International, Inc.
- Powell, R. R. (2006, Summer). Evaluation Research: An Overview. *Library Trends*, 55 (1), 102-120.
- United Way of America. (1996). Measuring program outcomes: A practical approach, 19th printing. Alexandria, VA: United Way of America.
- Unrau, Y. A., Garbor, P. A., & Gainell, Jr., R. M. (2007). Evaluation in social work: The art and science of practice (4th ed.). New York: Oxford University Press.
- Westerfelt, A., & Dietz, T. J. (2005). Planning and conducting agency-based research (3rd ed.). Boston: Pearson Education.
- Vivian R. Bergel, Ph.D., LSW, earned her MSW from West Virginia University and her Ph.D. from the University of Maryland at Baltimore. She is an associate professor and director of field instruction in the social work department at Elizabethtown College, Elizabethtown, PA. She was formerly chair of the department, having begun her career in teaching there in 1979. She has experience in clinical and geriatric social work.
- Peggy McFarland, Ph.D., LSW, earned her MSW from Marywood College and her Ph.D. from the University of Maryland at Baltimore. She is currently chair of the social work department at Elizabethtown College, where she has taught since 1990. Dr. McFarland maintains a private geriatric care management practice, Senior Management Services, which she co-founded in 1990. She has experience in adult day care, home health care, and dementia care. She has published in the *Journal of Gerontological Social Work*, *Journal of Social Work Education*, and *American Journal of Alzheimer's Disease*.