An assessment of goal-free evaluation: Case studies of four goal-free evaluations

Brandon W. Youker *, Allyssa Ingraham, Nicholas Bayer

Grand Valley State University, United States

ABSTRACT

This article provides a descriptive review of four goal-free program evaluations (GFE). GFE is an evaluation model where the independent evaluator is intentionally screened from the program's stated goals and objectives in hopes of reducing potential goal-related tunnel vision. The findings from these GFE case studies are focused in three areas: (1) elements of the programs that were evaluated and their existing evaluation contexts (e.g., pre-evaluation conditions, size of evaluation budget), (2) design of the GFEs (e.g., screening method, data collection methods), and (3) expertise of the goal-free evaluators (e.g., training, degrees attained). The findings indicate that, when employed, GFE is used as a qualitative data collection method; and the GFEs conducted have been relatively small in size and scope. The conclusions are that a more explicit operationalization of GFE is needed for increased use, and that systematic and empirical study comparing GFE with other evaluation models is warranted.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Goal-free evaluation (GFE) refers to an evaluation approach in which an independent external evaluator intentionally avoids knowledge of and reference to the program's stated or official goals and objectives. To reduce the potential for goal contamination, a liaison is appointed who intercepts all evaluation-related communiqués and materials to eliminate goal-oriented information to screen it from the evaluator. According to Scriven (1991), the logic behind avoiding stated goals and objectives has to do with:

> Finding out what the program is actually *doing* without being cued as to what it is *trying* to do. If the program is achieving its stated goals and objectives, then these achievements should show up; if not, it is argued, they are irrelevant. (180)

Scriven (1973) analogizes GFE to the double-blind pharmaceutical study as the goal-free evaluator, like the pharmaceutical evaluator, does not need to know the direction of the intended effect or the intended extent of the outcomes. Rather, the evaluator searches for program outcomes and works backward to determine if these effects were actually caused by the program (Youker, 2005a). Scriven (1991) suggests merit determination is accomplished by comparing the program's outcomes to the meeting of the consumers' relevant needs and thus, for Scriven, a consumer needs assessment is linked to GFE.

Forty years have passed since Scriven (1973) introduced GFE; and a multitude of evaluation scholars have recognized it as one of several evaluation models that should be considered for inclusion in the evaluator's toolbox (e.g., Davidson, 2005; Fitzpatrick, Sanders, & Worthen, 2004; Irving, 1979; Scriven, 1973; Stufflebeam, 2001; Worthen, 1990; Youker, 2005b). However, only a handful of evaluators claim to have conducted a goal-free program evaluation such as Belanger (2006), Berkshire, Kouame, and Richardson (2009), Gustufson (O. Gustufson, personal communication, April 27, 2006), House and Hogben (in Evers, 1980), James and Roffe (2000), Manfredi (2003), Matsunaga and Enos (1997), Scriven (Salasin, 1974), Stufflebeam (2001), Thiagarajan (1975), Welch (1976,1978), and Youker (2013). Yet, there is little information about whether, how, and when to design and implement GFE considering resource constraints. On the contrary, the literature consists of prescriptive claims usually regarding philosophical and theoretical arguments for or against GFE (Youker, 2013). The only practical recommendations for conducting GFE found in the literature suggest that one should attempt to observe and measure all relevant actual outcomes, effects, or impacts, intended or unintended, without being cued to the program's goals and objectives (Youker & Ingraham, 2013); to

* Corresponding author at: School of Social Work, Grand Valley State University, 353C DeVos Center, 401 West Fulton Street, Grand Rapids, MI 49504-6431, United States. Tel.: +1 616 331 6585; mobile: +1 269 744 4580.
E-mail addresses: youkerb@gvsu.edu (B.W. Youker), ingrahama@mail.gvsu.edu (A. Ingraham).

conduct a needs assessment; and to appoint an impartial goal screener to shield the evaluator from the goals (Scriven, 1974, 1991). In a recent attempt to further articulate general principles for guiding the goal-free evaluator, Youker (2013) proposed the following principles:

1. Identify relevant effects to examine without referencing goals and objectives.
2. Identify what occurred without the prompting of goals and objectives.
3. Determine if what occurred can logically be attributed to the program or intervention.
4. Determine the degree to which the effects are positive, negative, or neutral. (p. 434)

It quickly becomes apparent that this literature provides insufficient information for an evaluator trying to learn how to conduct a GFE. An additional challenge is finding cases of GFE in practice for extrapolating themes and developing further principles. Because of the scarcity of GFE use, this study examines four GFE reports. The hope is that lessons learned from these evaluations and reports may make future GFEs more effective and useful. For purposes of full disclosure, the author of this article was an evaluator for one of the programs being reviewed.

2. Methods

Four technical evaluation reports were reviewed. The evaluations were identified through questioning established evaluation scholars. This examination is limited in that these four evaluations in no way represent all GFEs conducted; nor do they represent all GFE technical reports ever written. They are however a sample of actual GFEs which were conducted in evaluating human service and educational programs. Some aspects of the evaluations were amenable to extraction from the evaluation reports while other aspects were not explicitly clear. Nevertheless, each evaluation report was analyzed, and the programs as well as their associated GFEs were compared.

To inform evaluation practice, this article details the commonalities found among the evaluation reports. These commonalities are categorized according to two evaluation characteristics conducive to being drawn from evaluation reports. The two

characteristics of program evaluation emphasized in this study are as follows:

- *Evaluation design.* An evaluation should follow a clear design in which the evaluation's method and plan are articulated.
- *Evaluator expertise.* Individuals trained and experienced in evaluation should be involved in the design and implementation of the evaluation.

By examining the two aforementioned evaluation characteristics, other equally important evaluation characteristics were not reviewed. For example, statistical approaches used for assessing, monitoring, and evaluating programs were not included in this review because none of the GFEs employed quantitative methods during data collection or analysis.

The findings are presented in four tables. First is an overview of the four programs evaluated; the second table is a review of the GFEs of these programs. These sections provide for a contextual understanding of the environment in which the GFEs existed. Following the overview of the programs and their evaluations is a description of the two evaluation characteristics of interest in this study, GFE design and goal-free evaluator expertise. The article concludes with a discussion of these GFEs as well as areas for further study.

3. Findings

3.1. Overview of goal-free evaluated programs included in evaluation assessment

Table 1 presents a general overview of the four programs that were evaluated via GFE and whose reports were included in this study. Knowing about these agencies and associated programs allows for an understanding of pre-evaluation conditions that influenced the decision to use GFE. This overview consists of (1) the type of program, (2) the partnering (or non-partnering) status of the agencies that administered the program, (3) the number of program sites, (4) the location of the program, (5) the dates and duration of program, and (6) the sponsor(s) of the program. Offering basic demographic information describes the agencies and associated programs which are the objects of this examination.

The type of program evaluated simply refers to what the programs are and do. Two of the evaluated programs were

Table 1
Overview of goal-free evaluated programs included in evaluation assessment.

Program evaluated	Program type/key interventions	Program partnership	Program sites	Program location	Program dates and duration	Program sponsor(s)
Making It Work Evaluation (MIW)	A welfare and housing/homeless prevention program	Yes, administered as a partnership between Kalamazoo County DHS and Housing Resources, Inc.	1	Kalamazoo County, Michigan	2001 to present	United Way and program partners
Kalamazoo Public Schools (KPS) Middle School Summer Enrichment Program (MSSEP)	A summer school for middle school students	No, primarily administered by KPS	1	Kalamazoo, Michigan	19 days of Summer 2005	No Child Left Behind Act of 2001: Title I Section 31a At Risk
Ke Aka Ho'ona on the Wai'anae Coast Community	A community self-help housing program	No, primarily administered by Families and Children for Empowerment and Development for Philippine Healthy Start	1	Oahu, Hawaii	1996–2000	Bulig Foundation and Consuelo Zobel Alger Foundation
Unnamed Program at Independent 4-Year Colleges	College-wide efficiency/cost reduction initiative	No, primarily independently administered by each college	2	Midwest USA	1974–1975	Hill Family Foundation

Notes: MIW, Making It Work; MSSEP, Middle School Summer Enrichment Program; Ke Aka Ho'ona, Ke Aka Ho'ona on the Wai'anae Coast Community; Colleges, Undisclosed 4-Year Colleges; KPS, Kalamazoo Public Schools; DHS, Department of Human Services.