

7-28-2020

Objectives Driven Participatory Evaluation Model

Dustin R. Saalman
Baker College, drsaalman@gmail.com

Follow this and additional works at: <https://digitalcommons.wayne.edu/jmasm>



Part of the [Applied Statistics Commons](#), [Social and Behavioral Sciences Commons](#), and the [Statistical Theory Commons](#)

Recommended Citation

Saalman, D. R. (2019). Objectives Driven Participatory Evaluation Model. *Journal of Modern Applied Statistical Methods*, 18(2), eP3474. <https://doi.org/10.22237/jmasm/1604190900>

This Quantitative Program Evaluation is brought to you for free and open access by the Open Access Journals at DigitalCommons@WayneState. It has been accepted for inclusion in *Journal of Modern Applied Statistical Methods* by an authorized editor of DigitalCommons@WayneState.

SPECIAL SECTION ON QUANTITATIVE PROGRAM EVALUATION

Objectives Driven Participatory Evaluation Model

Dustin R. Saalman

Baker College
Auburn Hills, MI

The ability to complete program evaluations of educational programming is typically restricted by the availability of resources, such as time, money and a trained evaluator. Although not a replacement for trained evaluators, promoting evaluative capacity and evaluative thinking within an organization can help mitigate this gap between talent and resources. Participatory evaluation is purported to increase organizational learning and promote evaluative thinking within an organization (Cousins & Whitmore, 1998). Objectives oriented evaluation is an easily understood evaluation method which provides a refined focus program outcome (Madaus & Stufflebeam, 1989). Using an internal evaluation of a new faculty onboarding course at a private non-profit college system, a mixed methods study was completed to explore the use of a participatory evaluation program evaluation with the use of the program objectives as an advanced organizer. An explanatory sequential design was employed utilizing quantitative findings to collect qualitative data to further explore the participants' experiences completing the program evaluation. This combined evaluation methodology met the criteria posited in Daigneault and Jacob (2009) and Toal (2009) to be considered participatory in its implementation. It also involved participants in ways which provided them experiences which helped develop evaluative thinking, skills, and beliefs.

Keywords: program evaluation, participatory evaluation, research on evaluation, mixed-methods

Introduction

The beginning of most program evaluation textbooks is a description of the importance of program evaluation and the need for all programs to be evaluated for

accountability and to demonstrate the merit and worth of the program (Fitzpatrick, Sanders, & Worthen, 2011; Stufflebeam & Shinkfield, 2007; Stufflebeam & Coryn, 2014). In the context of education, educational programming represents a broad category of activities, for both instructors and learners, which take place in and around the educational environment, and not all programs come with a funding source to support formal evaluations of program outcomes. Nonetheless, knowing the outcomes of these programs is important to stakeholders such as educational administrators, boards, and policy makers who are accountable for funding and organizational effectiveness. Evidence of program outcomes is also important for improving programs to better meet the needs of the targeted audience. However, if the ability to complete formal evaluations of educational program is potentially limited by the availability of funding to pay for the services of an external evaluator or having availability of an internal evaluator, there is a gap between the stated importance of completing program evaluations and the availability of resources to complete them. One way of mitigating this limitation of resources is to build the evaluative capacity of the organization by providing opportunities for non-evaluative stakeholders to have training and experiences in program evaluation thereby increasing the capacity for evaluative thinking within the organization. This can be achieved by engaging diverse sets of stakeholders in the evaluation process and using an evaluation method that is eclectic in both its design and methodology.

Evaluation Models

Objectives oriented evaluation approaches are the most prevalent when evaluating programs due to a high degree of simplicity and clarity when reporting findings to stakeholders. These forms of evaluation are applicable when there are focused programs with supportable objectives (Stufflebeam, 2001). This describes most educational programs as goals and objectives are a fundamental building block of most program design (Richlin, 2006) and often a required element when applying for funding sources (Karsh & Fox, 2014). In objectives-oriented evaluations the evaluation criteria are defined by the program objectives (Fitzpatrick, Sanders, & Worthen, 2011; Stufflebeam, 2001). Due to limitations in time and resources, these sorts of evaluations have limited stakeholder involvement and operate within the limited scope of the program objectives (Fitzpatrick, Sanders, & Worthen, 2011).

Different evaluation models were developed to provide opportunity to define broader stakeholder groups and allow for more expanded opportunities for stakeholder involvement in the actual evaluation. These participant-oriented evaluation models were developed in some cases to create opportunities to better

OBJECTIVES DRIVEN PARTICIPATORY EVALUATION MODEL

capture the context of the program and better promote use of evaluation results, such as practical-participant evaluation and stakeholder evaluations (Stufflebeam, & Coryn, 2014). Whereas others were developed to promote social justice values, such as transformative-participatory evaluation and empowerment evaluation (Stufflebeam, & Coryn, 2014). Both have the value of promoting plurality of perspectives in the development and implementation of the evaluation.

However, participant-oriented models are not without their criticisms. Stufflebeam (2001) described these approaches (referred to as constructivist evaluation) as being “heavily philosophical, service oriented, and paradigm driven” (pg. 71). They are opposed to the positivist, deterministic, and reductionist structure and are an attempt to make sense of a variety of constructions emerging from the variety of stakeholders, avoiding the idea one of the constructions is “true” (Stufflebeam, 2001, p. 73). They also present barriers when stakeholders at higher levels within the organization are uncomfortable with the processes and outcomes of these sorts of evaluation models (Stufflebeam, 2001). Daigneault and Jacob (2009) described other shortcomings of these approaches as lacking satisfactory conceptualization and operationalization.

Statement of the Problem

An evaluation design which promotes the inclusion of a diverse set of stakeholders and includes training and experience completing a program evaluation and borrows the use of program objectives from the objective-oriented evaluation methods as an organizational framework to guide the development of the evaluation framework has the potential to capitalize on the strengths of both methods while mitigating many of the stated limitations or concerns of both methods. However, limiting the scope of the evaluation to the program objectives risks whether or not the evaluation can truly be considered participatory and therefore is a risk to capitalizing on the purported benefits of utilizing a participatory evaluation model.

Therefore, the research question addressed in this study is: When limiting the scope of a participatory evaluation using an objectives-oriented evaluation model as an advanced organizer, does the evaluation meet the criteria provided by Daigneault and Jacob (2009) using the Participatory Evaluation Measurement Instrument (PEMI) and Evaluation Involvement Scale (EIS) provided by Toal (2009)?

Review of Literature

Program evaluation formalized as a field of research and investigation in the United States and Canada in the 1960s as the United States government began investing in large educational programs that required evaluations as part of government accountability. As program evaluation grew and became a more common practice in education, healthcare, and other non-profit organizations the definition of program evaluation has been redefined. However, Stufflebeam and Coryn (2014) defined evaluation as “the systematic assessment of an object’s merit, worth, probity, feasibility, safety, significance, and/or equity” (p. 7), which is an extension of the Joint Committee’s 1994 definition of evaluation.

As a practice, program evaluation, is broad and varied. There are twenty-three different approaches to program evaluation (Stufflebeam, 2001; Stufflebeam & Coryn, 2014) which were categorized by Stufflebeam and Coryn (2014) as pseudo-evaluation approaches, improvement and accountability-oriented approaches, social agenda and advocacy approaches, and eclectic approaches. Although the approaches to program evaluation are varied in primary goals, advanced organizers, sources of the questions addressed, and methods typically employed, there are general functions and steps in evaluation design which are common regardless of the approach, impetus, or context for the evaluation (Stufflebeam & Coryn, 2014). Stufflebeam and Shinkfield (2007) proposed four functions common to conducting any evaluation: 1) information collection, 2) organization, 3) analysis, and 4) reporting.

Although the stated functions are common in the design of the evaluation there are a multitude of ways that the activities in the evaluation can be carried out depending on the specific evaluation methodology. This is determined in large part by the political context in which the evaluation is being carried out and the philosophical assumptions and theories on which the evaluation is based (Mertens, 2018).

Objectives-Oriented Evaluation

Objectives-oriented evaluation was classified by Stufflebeam (2001) as a quasi-evaluation model which uses the objectives of an educational program as an advanced organizer to determine the extent to which the program’s objectives are being met. Early versions of objectives-oriented evaluation models were developed by Ralph Tyler (Madaus & Stufflebeam, 1989) and Malcom Provus (Fitzpatrick, Sanders, & Worthen, 2011). Objectives oriented evaluation is situated in the methods focused evaluation theories (and the theorists represented in this group of

OBJECTIVES DRIVEN PARTICIPATORY EVALUATION MODEL

evaluation models are rooted in the positivist/postpositivist paradigms. As a result, importance is placed on objectivity and quantitative methods are commonly associated with these sorts of evaluations (Alkin, 2004; Mertens, 2018).

As one of the earliest models of objectives-orientated evaluation, The Tylerian (Tyler, 1931) evaluation approach was developed in the 1930s. Ralph Tyler developed this method when contracted to work with undergraduate biology faculty at the Ohio State University. The goal was to develop achievement tests as a means of tracking students to help reduce failures and dropouts. Using course learning objectives, Tyler evaluated the undergraduate biology courses. (Tyler 1931; Wraga, 2017). The discrepancy between what the student ought to have learned and what the student demonstrated as learned provided an evaluation of the course curriculum.

Tyler (1931) advocated the use of multiple indicators to measure outcomes, including direct and indirect measures; Tyler (1931) opined evaluators should employ a broad array of data collection techniques to support the findings of the evaluation. To that end, the sources of evidence went beyond test scores and included observations, self-reporting instruments, interviews, and examples of student products (Madaus & Stufflebeam, 1989). This flexibility in the model laid foundations for additional objectives-oriented evaluation approaches to be designed.

The creation of the Discrepancy Evaluation Model (DEM) was credited to both Malcom Provus and Andres Steinmetz (Fitzpatrick, Sanders, & Worthen, 2011; Madaus, Scriven, & Stufflebeam, 1983). Provus' DEM is an objectives-oriented approach to evaluation, in the Tylerian tradition, which uses similar design but is used outside of the curriculum evaluation context by developing standards to use as comparison. This model was originally developed by Provus (Worthen & Sanders, 1973) as means of providing information to the administration in the Pittsburgh Public School system as a way to support the administration's decision making.

According to Madaus, Scriven, and Stufflebeam (1983), Andres Steinmetz contributed to DEM by providing a scheme with which to respond to the task of evaluating educational programs. Steinmetz (1983) described the goal of DEM as "to make judgements about the worth or adequacy of an object based on the discrepancy information between the standard and the performance" (p. 80) and as a result provided the formula $S - P = D$, where S is the standard, P is the performance, and D is the discrepancy. The role of the evaluator in a DEM evaluation is to work with the client to determine the standards the performance will be compared to and to seek out technical expertise to measure the performance.

Once the performance data is collected the evaluator works with the client to compare the performance to the standards and determine the discrepancy.

The strengths of objectives-oriented evaluation approaches frequently cited are related to its ease of use and narrow focus, and the methods are easily understood and easy to follow (Fitzpatrick, Sanders, & Worthen, 2011). This makes it a desirable model for evaluations completed internally and appropriate for programs and projects which “have clear, supportable objectives” (Stufflebeam, 2001, p. 16). Completing the evaluation requires that stakeholders involved in the evaluation reflect on the program and its intentions. The sources of data, which inform the outcomes of the program, need to be dependent upon the nature of the objective (Madaus & Stufflebeam, 1989), which allows for broad considerations of both qualitative and quantitative data collection and analysis.

The role of the evaluator in objectives-oriented evaluation models focuses on working with subject matter experts, program directors, or other program stakeholders to define the standards or objectives the program is designed to achieve and developing the means to compare them with the actual program outcomes (Fitzpatrick, Sanders, & Worthen, 2011; Madaus, Scriven, & Stufflebeam, 1983; Wraga, 2017; Madaus & Stufflebeam, 1989). The evaluator is not establishing the objectives or the standards of the program or judging any discrepancy resulting from the evaluation. This is left to the stakeholders who helped establish the program objectives or standards as part of the evaluation. However, this introduces an inherent bias as a limitation of these types of evaluation methods (Steinmetz, 1983). However, this limitation can be mitigated depending on how broad the stakeholder group is that provides input on the objectives and any determined discrepancy.

Additional limitations of objectives-oriented evaluation approaches often cited are the limitations imposed by the single-minded focus of the evaluation on the objectives or standards of a program or course (Fitzpatrick, Sanders, & Worthen, 2011). These self-imposed blinders, critics claim, can cause evaluators to ignore other program outcomes or make incomplete judgments about the program. As a result, the outcomes of the evaluation can be seen as leading to terminal information that is neither timely nor pertinent to help advance or improve the program (Stufflebeam, 2001).

Participatory Evaluation

Participatory oriented evaluation models were first developed in the 1970s, out of growing concerns evaluators did not really understand the phenomena which they

OBJECTIVES DRIVEN PARTICIPATORY EVALUATION MODEL

were evaluating (Fitzpatrick, Sanders & Worthen, 2011). Emerging models stressed “experience with program activities and settings as well as involvement of program participants, staff, and managers in the program evaluation” (Fitzpatrick, Sanders, & Worthen, 2011, p. 190). Cousins (2005) described participatory evaluation as “members of the evaluation, community and members of the stakeholder groups, relative to the evaluand each participate in some or in all of the shaping and/or technical activities required to produce evaluation knowledge leading to judgments of merit and worth and support for program decision making.” This type of evaluation method is representative of the use, values, and social justice evaluation theories (Alkin, 2004).

Participatory evaluation as an evaluation model, as opposed to the category of models, grew out of the stakeholder-based evaluation model. It was an early attempt to move toward evaluation results which were more responsive to the needs of the users of the results (Cousins & Earl, 1992). In the stakeholder-based model the evaluator serves as the principal investigator and stakeholders are used in a consultative fashion; whereas, in a participatory evaluation model the evaluator is seen as a technical expert and facilitator, and the stakeholders largely take on the responsibility of the evaluation activities (Cousins & Earl, 1992). Cousins and Whitmore (1998) delineated participatory evaluation into two distinct streams: Practical-Participatory Evaluation (PPE) and Transformative-Participatory Evaluation (TPE). Transformative-Participatory Evaluation endeavors to invoke participatory principles and actions in order to democratize social change (Cousins & Whitmore, 1998). TPE is typically used when completing evaluations of programs intended to benefit marginalized communities and came out of evaluation work in developing countries. The intent is to empower the members of those communities to benefit from the evaluation process and results (Mertens, 2018). Whereas, Practical-Participatory is focused on organizational learning and capacity building (Mertens, 2018). For the purposes of this study Transformative-Participatory Evaluation is not an appropriate evaluation model. Henceforth, participatory evaluation will be in reference to the Practical-Participatory Evaluation stream. Increasing evaluation utilization is one of the primary benefits of participatory evaluation.

Criticisms of participatory evaluation models focused on the increased need for resources to complete the evaluation and the credibility of the evaluation to those who are external to the evaluation itself (Fitzpatrick, Sanders, & Worthen, 2011). Having those who are closest to the program completing the evaluation increases the risk of inserting bias into the process. Daigneault and Jacob (2009)

also identified ambiguous conceptualization of participatory evaluation as a concern.

The implementation of a participatory evaluation will vary based on the context and program being evaluated. However, the common element which need to be present for a participatory evaluation to be carried out is partnership between trained program evaluators and program stakeholders, as Moreau (2017) noted, “More specifically, the program evaluators (who have technical expertise in evaluation theory and methods) collaborate with program stakeholders (who have an understanding of a program and how it is supposed to operate) to evaluate a program” (p. 334). The stakeholders should represent not only those responsible for the development and delivery of a program but also those who are intended to participate and benefit from the program. Evaluation team size and makeup of stakeholders will vary depending on the scope of the evaluation and the size of the program.

Stakeholder involvement in completing the evaluation is one of the hallmarks of participatory evaluation, with the intent to promote evaluation use, evaluative thinking, and evaluative capacity. Stakeholders are “actively involved in the evaluation design, data collection, analyses, interpretation, and dissemination efforts” (Moreau, 2017, p. 334). Participatory evaluations should begin with the training of the stakeholders who will make up the evaluation team to provide them with the necessary knowledge and skills to successfully complete the various evaluation tasks. Training should be a first step in a participatory evaluation and should continue to occur throughout the evaluation as needs and skill gaps present themselves.

Following training, the stakeholders begin the evaluation by developing an evaluation framework (Moreau, 2017). The evaluation framework provides an overview of the major evaluation questions and establishes what will be evaluated, who will be involved in the evaluation activity when the activities will take place, what evaluation and analysis methods will be used, and how the findings will be consolidated and the results shared (Canadian International Development Agency, 2001). The framework developed is unique to each evaluation. The process will vary depending on the context of the evaluation and the makeup of the evaluation team.

Once the evaluation framework is determined the evaluation can be carried out by the team. All participants in the evaluation should be completing the evaluation activities with support from the trained evaluator. Participatory evaluation is method agnostic, so the evaluator should be prepared to support

participants' analysis of data using qualitative, quantitative, and mix-methods (Moreau, 2017).

Defining and Measuring Participation

To some extent all evaluations include participation. The evaluator must interact with various stakeholder groups to develop and complete the evaluation (King, Cousins & Whitmore, 2007). Determining what type of participation in an evaluation qualifies as participatory evaluation is an ambiguity which theorists have struggled with. Cousins and Whitmore (1998) set forth a framework for participatory evaluation which identified three distinct process dimensions for participatory inquiry: a) Control of the evaluation process, b) Stakeholder selection for participation, and c) Depth of participation. Each dimension is independent of the other dimensions and exists on a spectrum of the extent to which the participation in the evaluation activities is limited or extensive.

Building from the conceptual framework provided in Cousins and Whitmore (1998), Daigneault and Jacob (2009) conceptualized and operationalized participatory evaluation by outlining the key decision points defined in participatory evaluation:

1. *Evaluation questions and issues definition/methodological design* – characterized as the moment when a decision is made about the framing of the evaluation including selection of evaluation questions, theoretical framework methods, techniques, and instruments.
2. *Data collection and analysis* - characterized by making decisions about how to concretely collect, assemble, code, and analyze data as well as carrying these tasks out.
3. *Judgements and recommendations formulation* – characterized by making decisions and determinations of merit and worth of a program and formulating suggestions for future action.
4. *Report and dissemination of evaluation findings* – characterized by the decisions about reporting and diffusion of evaluation findings and implication (Daigneault & Jacob, 2009, p. 339).

Each decision point is considered a dichotomous indicator of participation and typed as involvement of nonevaluative stakeholders in the task (presence of the indicator) or no involvement of nonevaluative stakeholders in the task (absence of the indicator) (Daigneault & Jacob, 2009, p. 339). Daigneault and Jacob (2009)

developed coding schemes, using the m of n rule, where $m = 1$ and $n = 4$, giving a weight of .25 to each indicator. These schemes are shown in Tables 1 and 2 for each of the dimensions of participatory inquiry posited by Cousins and Whitmore (1998) and are used to establish cut-off points for determining whether an evaluation can be considered participatory. The threshold for whether or not an evaluation could be considered participatory, according to Daigneault and Jacob (2009), was the presence of all three fundamental attributes of participation, and each category must have at least a score of .25 or greater. Daigneault and Jacob (2013) revised the process by adding a final step of averaging the scores across the three dimensions as an indicator a measurement of overall participation.

Table 1. Coding Scheme for Extent of Involvement, Diversity (Daigneault and Jacob, 2009, p. 340)

Number of Tasks Nonevaluative Stakeholders are Involved In	Level of Membership	
	Intuitive Label	Numerical
0	None	0.00
1	Limited/weak	0.25
2	Moderate	0.50
3	Substantial/strong	0.75
4	Full	1.00

Table 2. Coding Scheme for Control of the Evaluation Process (Daigneault and Jacob, 2009, p. 340)

Level of Membership	
Intuitive Label	Numerical
Exclusive control by evaluator and/or nonparticipating evaluation sponsor	0.00
Limited/weak control by nonevaluative participants	0.25
Shared control between nonevaluative participants and evaluator and/or nonparticipating evaluation sponsor	0.50
Substantial/strong control by nonevaluative participants	0.75
Exclusive control by nonevaluative participants	1.00

The Evaluation Involvement Scale (EIS) (Toal, 2009) was developed as a response to a call for instruments designed to measure the level of involvement of

OBJECTIVES DRIVEN PARTICIPATORY EVALUATION MODEL

stakeholders in a participatory evaluation in multisite settings. Due to the limitations of the program design the EIS only measures the third dimension of participatory inquiry posited in Cousins and Whitmore (1998), depth of participation. The scale was developed using key evaluation decision points posited in Burke (1998) as a general framework. Each key decision point was assigned to the three evaluation stages: 1) evaluation planning, 2) implementation, and 3) communication of results, as well as additional items the development team added to provide clarity or distinction to the decision points, resulting in thirteen item scale shown in Table 3 (Toal, 2009). Participants respond to each item rating the extent to which they were involved in the different evaluation activities using a 1:4 scale: 1 = No; 2 = Yes, a little; 3 = Yes, some; 4 = Yes, extensively (Toal, 2009, p. 354).

Table 3. Evaluation Involvement Scale Items (Toal, 2009, p. 356)

-
1. Discussions that focused the evaluation
 2. Identifying evaluation planning team members
 3. Developing the evaluation plan
 4. Developing data collection instruments
 5. Developing data collection processes
 6. Collecting data
 7. Reviewing collected data for accuracy and/or completeness
 8. Analyzing data
 9. Interpreting collected data
 10. Writing evaluation reports
 11. Reviewing evaluation reports for accuracy and/or completeness
 12. Presenting evaluation findings (e.g., to staff, to stakeholders, to an external audience)
 13. Developing future project plans based on evaluation results
-

Methodology

Design

The purpose of this study is to explore, when limiting the scope of a participatory evaluation using an objectives-oriented evaluation model as an advanced organizer, whether or not the evaluation meet the criteria provided by Daigneault and Jacob

(2009) using the Participatory Evaluation Measurement Instrument (PEMI) and Evaluation Involvement Scale (EIS) provided by Toal (2009) to be considered participatory.

The mixed methods research design employed in the study was an explanatory sequential design, which consisted of two distinct interactive phases illustrated in Figure 1 (Creswell & Plano Clark, 2018). The notation for the design was QUAN → qual indicating there were two strands in the sequence, the quantitative method occurring first and receiving the greater emphasis in addressing the study's purpose. The quantitative method is then followed by the qualitative methods to help explain the quantitative results (Creswell & Plano Clark, 2018). Approval by the Institutional Review Board (IRB) was received for both phases.

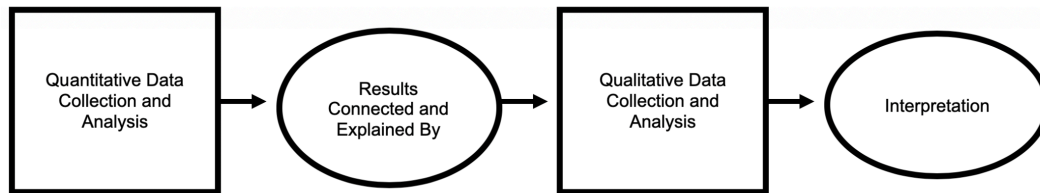


Figure 1. Explanatory Sequential Design

Setting Description

The setting for the study was a private non-profit college system located in the midwestern United States. The college consists of seven satellite campuses and one online campus. The college is accredited by the Higher Learning Commission and offers programs that award certificates, associates, bachelors, masters, and doctoral degrees (HLC, 2019). A team of stakeholders from the Academics division of the college was organized to perform an internal evaluation of the new faculty onboarding program referred to as the First Year Faculty Experience (FYFE).

The evaluation team members completed a self-paced program evaluation tutorial based on the Framework for Program Evaluation available from the Centers for Disease Control (2018). The tutorial was delivered via the institution's learning management system prior to the initial planning meeting. The team then participated in the development and implementation of the program evaluation.

The stakeholders designing and implementing the program evaluation provided a convenience sample for the study. The stakeholder groups represented on the team were campus and system leadership, administrative staff, program

OBJECTIVES DRIVEN PARTICIPATORY EVALUATION MODEL

officials who have faculty from their program complete the FYFE prior to teaching, and adjunct faculty who completed the FYFE. Stakeholder types are categorized as traditional and non-tradition in Table 4. This sampling plan was applied to Phase I and Phase II of the mixed-methods design.

Table 4. Traditional and Non-Traditional Stakeholder Types

Stakeholder	Traditional	Non-Traditional
Program Leadership	X	
Program Staff	X	
Academic Program Directors		X
Adjunct faculty consuming the program		X
Students of the adjunct faculty		X

Instrument

The quantitative data was collected using two instruments, the Participatory Evaluation Measurement Instrument (PEMI) and the Evaluation Involvement Scale (EIS) (Daigneault & Jacob, 2009; Toal, 2009). The PEMI is a nonnormative instrument, which measures the evaluation participants perception of participation on the three dimensions of Participatory Inquiry outlined by Cousins and Whitmore (1998) (Daigneault & Jacob, 2009; Daigneault & Jacob, 2013). The instrument was validated using the Instrument Development and Construct Validation process developed by Onwuegbuzie, Bustamante, and Nelson (2010), a 10-phase mixed-methods validation process (Daigneault & Jacob, 2013). Each dimension in the instrument has the same scale with different intuitive labels for the respective domain, 0 = .00, 1 = .25, 2 = .50, 3 = .75, 4 = 1.00. For an evaluation to be considered participatory it must receive at least an average of .25 on each domain. Finally, the scores from the three domains are averaged to measure overall participation (Daigneault & Jacob, 2013). Therefore, this instrument addresses the question of whether or not an evaluation can be considered participatory.

The second instrument used to collect quantitative data was the EIS. Using Messick's unitary concept of validity as a framework, the EIS was validated using data collected from a multi-site evaluation, with a resulting α of .94, suggesting high internal consistency (Toal, 2009). The scale was developed using the evaluation key decision points posited in Burke (1998) as a general framework. Each key decision point was assigned to the three evaluation stages: 1) evaluation planning, 2) implementation, and 3) communication of results, as well as additional

items the development team added to provide clarity or distinction to the decision points, resulting in thirteen item scale shown in Table 3 (Toal, 2009). Participants responded to each item rating the extent to which they were involved in the different evaluation activities using a 1:4 scale: 1 = No; 2 = Yes, a little; 3 = Yes, some; 4 = Yes, extensively, or “I don't think this activity took place” (Toal, 2009, p. 354). The EIS addresses the question of how participatory the participatory evaluation was.

Following the completion of the program evaluation, a survey was distributed to the participants via email Qualtrics (vers. XM) comprised of questions from the PEMI and the EIS. The results collected from the surveys were uploaded into an SPSS database. A separate form of analysis was completed on each data set. All statistical analyses were conducted at the nominal alpha level = 0.05.

Participatory Evaluation Measurement Instrument (PEMI) Descriptive statistics, including mean and standard deviations, were generated for each of the three domains of the PEMI. The Kruskal-Wallis test (Wilcox, 2011) was used to determine if there are differences in mean responses between the three domains represented in the PEMI.

$$H_0 : F_{(\text{Extent of Involvement})} = F_{(\text{Diversity of Participants})} = F_{(\text{Control of the Evaluation Process})}$$

$$H_1 : F_{(\text{Extent of Involvement})} \neq F_{(\text{Diversity of Participants})} \neq F_{(\text{Control of the Evaluation Process})}$$

The results of the one-way analysis of variance provided evidence of differences between the three categories of the PEMI and was used for further exploration using qualitative methods.

Evaluation Involvement Scale (EIS) For each question on the EIS descriptive statistics, including mean and standard deviations, were generated for all question responses on the EIS. Additionally, the Sign Test was completed to calculate the median value of each question for the population. For a sample size $n = 10$ the critical region is S1 (Neave & Worthington, 1988).

$$H_0 : = 3$$

$$H_1 : \neq 3$$

Phase II: Qualitative

Interview Protocol Development The interview protocol was developed to complete semi-structured interviews with the members of the evaluation planning committee to explain their responses on the two instruments and their responses to the survey on evaluative thinking. The analysis of the quantitative data was used to identify which results are significant and needed to be further explained through qualitative data collection (Creswell & Plano Clark, 2018). The interviews with the research participants were conducted virtually via WebEx. The interviews were recorded and transcribed by the researcher.

Data Analysis

A general inductive analysis approach was used for analyzing the data collected from the semi-structured interviews and the evaluation planning sessions. Thomas (2006) posited “the primary purpose of the inductive approach is to allow research findings to emerge from the frequent, dominant, or significant themes inherent in raw data, without the restraints imposed by structured methodologies” (p. 238).

Results

Phase I: Quantitative

Participatory Evaluation Measurement Instrument Following the completion of the program evaluation, participants completed the Participatory Evaluation Measurement Instrument (PEMI) (Daigneault & Jacob, 2009) via Qualtrics (ver. XM). All participants ($N = 9$) completed the PEMI, rating the extent of their involvement in the First Year Faculty Experience (FYFE) program evaluation in the Extent of Involvement, Diversity of Participants, and Control of the Evaluation Process domains. The responses were converted to a numerical scale ranging from .00 to 1.00. SPSS (ver. 25) was used to generate descriptive statistics and complete analysis of the data. Cronbach’s alpha was calculated to assess the reliability of the PEMI. The instrument was found to have acceptable reliability (3 items; $\alpha = .76$).

The descriptive statics for each item on the instrument are provided in Table 5. The mean response for the Extent of Involvement domain ($M = .6944$, $SD = .1667$) indicates participants opined the extent of involvement was between moderate involvement and substantial/strong involvement. The mean response for the Diversity of Participants domain ($M = .6667$, $SD = .2165$) indicates the

diversity of the evaluation participants was between moderate diversity and substantial/strong diversity. The mean response for the Control of the Evaluation Process domain ($M = .5278$, $SD = .2319$) indicates the control of the evaluation process was shared between the nonevaluative participants and the evaluator.

Table 5. PEMI Descriptive Statistics and Mean Ranks

Domain	N	M	SD	Min	Max	M Rank
Involvement	9	0.6944	0.1667	0.5	1.00	16.39
Diversity	9	0.6667	0.2165	0.5	1.00	14.56
Control	9	0.5278	0.2319	0.0	0.75	11.06

The Kruskal-Wallis test was used to determine if there were statistically significant differences in mean responses between the three domains represented in the PEMI. All analyses here and below were completed with nominal $\alpha = 0.05$. The results Kruskal-Wallis Test are listed in Table 6. There was not a statistically significant difference between the mean rating for the three domains of the PEMI ($H(2) = 2.506$, $p = .286$) with a mean rank of 16.39 for the Extent of Involvement domain, 14.56 for the Diversity of Participants domain, and 11.06 for the Control of the Evaluation Process domain.

Table 6. Kruskal-Wallis Test Statistics

Kruskal-Wallis H	2.506
df	2.000
Asymp. Sig.	0.286

Based on the results of the analysis of the PEMI the following questions were generated and added to the semi-structured interview protocol for the qualitative phase of the study to further explain the responses that participants provided:

1. Reflecting on the level of membership in the evaluation, what were the considerations for you when selecting your response?
2. Reflecting on the diversity of participants in the evaluation, what were the considerations for you when selecting your response?

OBJECTIVES DRIVEN PARTICIPATORY EVALUATION MODEL

3. Reflecting on the control over the evaluation process during the evaluation, what were the considerations for you when selecting your response?

Evaluation Involvement Scale Following the completion of the program evaluation, participants completed the Evaluation Involvement Scale (EIS) (Toal, 2009) via Qualtrics (ver. XM) All participants ($N = 9$) completed the EIS indicating the extent they participated in the thirteen evaluation activities listed in the instrument, on a scale of 1 = No involvement to 4 = Yes, extensive involvement (Toal, 2009). The descriptive statistics for each item on the instrument are provided in Table 7. Cronbach's alpha was calculated to assess the reliability of the EIS. The instrument was found to have good reliability (13 items; $\alpha = .83$).

Table 7. EIS Descriptive Statistics and Levels of Significance

Item	N	M	SD	Min	Max	Sig.
EIS1	9	3.11	0.601	2	4	0.564
EIS2	9	1.56	0.726	1	3	0.009
EIS3	9	2.33	0.866	1	3	0.063
EIS4	9	2.22	0.833	1	3	0.038
EIS5	9	2.67	0.707	1	3	0.180
EIS6	9	2.67	1.000	1	4	0.276
EIS7	9	2.78	0.441	2	3	0.157
EIS8	9	2.89	0.782	2	4	0.655
EIS9	9	2.78	1.093	1	4	0.458
EIS10	9	1.67	1.000	1	4	0.015
EIS11	9	1.56	1.014	1	4	0.012
EIS12	9	1.22	0.667	1	3	0.005
EIS13	9	1.89	1.167	1	4	0.028

Although the Sign test is not available in SPSS, the one sample Wilcoxon Signed-Rank Test, a functional equivalent, was used to analyze the results of the EIS to identify which activities on the instrument resulted with ratings that were statistically significant (Neave & Worthington, 1988). The hypothesized median rating for each activity was three. Six activities had ratings that were statistically significant: 1) EIS2: Identifying evaluation planning team members ($M = 1.56$, $SD = .726$, $p = .009$), 2) EIS4: Developing data collection instruments ($M = 2.22$, $SD = .833$, $p = .038$), 3) EIS10: Writing evaluation reports ($M = 1.67$, $SD = 1.00$, $p = .015$), 4) EIS11: Reviewing evaluation reports for accuracy and/or completeness

($M = 1.56$, $SD = 1.014$, $p = .012$), 5) EIS12: Presenting evaluation findings ($M = 1.22$, $SD = .667$, $p = .005$), and 6) EIS13: Developing future project plans based on evaluation results ($M = 1.89$, $SD = 1.167$, $p = .028$).

A review of the individual submissions for the items with significant results demonstrates that these activities were rated at higher levels than the other activities as “No involvement” or “Yes, a little involvement,” by the participants. A sum of these two ratings for each item on the instrument is provided in Table 8.

Table 8. EIS Items and Sum of "No" or "Yes, a little" Ratings

Item	SUM
EIS1	1
EIS2	8
EIS3	4
EIS4	5
EIS5	2
EIS6	2
EIS7	2
EIS8	3
EIS9	2
EIS10	8
EIS11	8
EIS12	8
EIS13	6

Based on the responses to the Evaluation Involvement Scale, the items with ratings that stood out as being significant were:

- Identifying evaluation planning team members
- Developing data collection instruments
- Writing evaluation reports
- Reviewing evaluation reports for accuracy and/or completeness
- Presenting evaluation findings (e.g., to staff, to stakeholder, to an external audience)
- Developing future project plans based on evaluation results
- Looking over your responses to the instrument, what were the factors you considered when providing your responses to these questions?

OBJECTIVES DRIVEN PARTICIPATORY EVALUATION MODEL

To determine the extent of involvement for each task included within the instrument the coding scheme posited by Toal (2009) is applied based on the mean rating:

- 1 = No Involvement
- 1.01 – 2.00 = Little Involvement
- 2.01 – 3.00 = Some Involvement
- 3.01 – 4.00 = Extensive Involvement (p. 355)

The coding scheme for each task is provided in Table 9 based on the responses from the study participants.

Table 9. Coding Scheme Applied to EIS Results

Item	Involvement Item	<i>M</i>	Involvement Level
EIS1	Discussions that focused the evaluation	3.11	Extensive
EIS2	Identifying evaluation planning team members	1.56	Little
EIS3	Developing the evaluation plan	2.33	Some
EIS4	Developing data collection instruments	2.22	Some
EIS5	Developing data collection processes	2.67	Some
EIS6	Collecting data	2.67	Some
EIS7	Reviewing collected data for accuracy and/or completeness	2.78	Some
EIS8	Analyzing data	2.89	Some
EIS9	Interpreting collected data	2.78	Some
EIS10	Writing evaluation reports	1.67	Little
EIS11	Reviewing evaluation reports for accuracy and/or completeness	1.56	Little
EIS12	Presenting evaluation findings (e.g., to staff, stakeholders, an external audience)	1.22	Little
EIS13	Developing future project plans based on evaluation results	1.89	Little
<i>Overall</i>		2.26	Some

Phase II: Qualitative

For Phase II, follow-up interviews were held individually with the members of the evaluation team after completing the program evaluation ($N = 9$). The interviews were conducted utilizing a semi-structured interview protocol comprised of the follow-up questions identified in Phase I as a result of the quantitative data analysis. The interviews were recorded and transcribed by the researcher.

A general inductive approach (Thomas, 2006) was employed for analyzing the transcripts. Three transcripts were initially analyzed to develop a code book with descriptions and examples for each theme that would then be used for coding of the remaining transcripts. The transcripts and code book were shared with two additional coders to validate the codes and demonstrate reliability of coding. Coder #1 results were used as confirmation of the researchers coding and Coder #2 results were used to identify additional follow-up coding for the researcher and Coder #1.

From the analysis five major themes were identified:

1. Control of the evaluation process,
2. Participant involvement,
3. Evaluation team diversity,
4. Participants' prior experience, and
5. Use

A graphical representation of the five themes and their subthemes is provided in Figure 2.

Control The control theme emerged as a discussion of the role that the trained evaluator fulfilled throughout the evaluation process and shared control with the participants. Two subthemes were identified through the analysis, 1) evaluator control and 2) participant control. Control was discussed in context of the selection of the members of the evaluation team, leading of the discussion, the role of the evaluator as guide versus decision maker, and assignment of evaluation tasks.

Participant Involvement The participant involvement theme emerged as a discussion of the various activities that the participants completed throughout the evaluation and the level of involvement in the process. The types of activities that were identified through the follow-up interviews were collecting data, reviewing data, collaborating with other members of the team to complete data collection, and providing feedback to other members of the team.

OBJECTIVES DRIVEN PARTICIPATORY EVALUATION MODEL

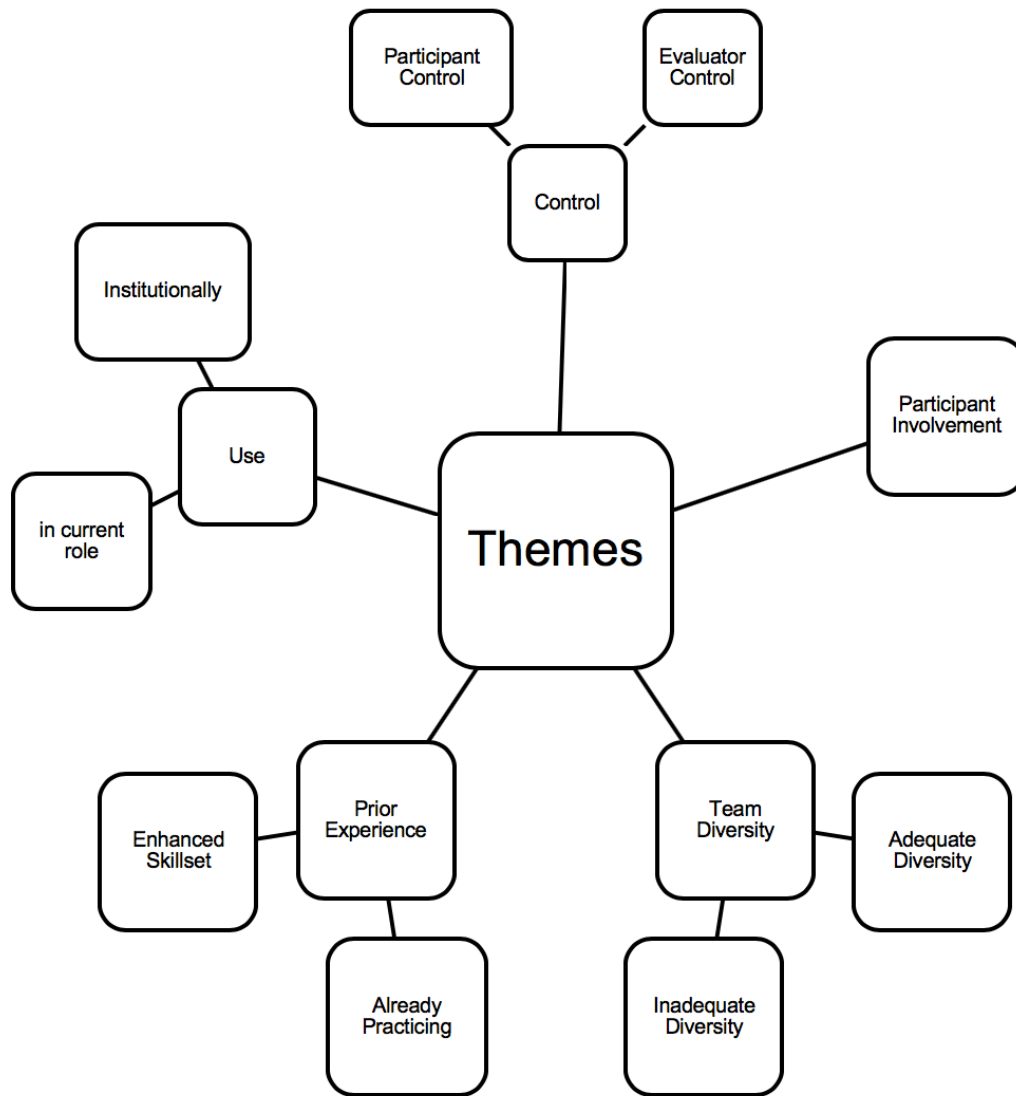


Figure 2. Major Themes and Subthemes Represented in Qualitative Data

Team Diversity

The team diversity theme emerged as a discussion about the adequacy of the diversity of the roles represented on the evaluation team as well as academic backgrounds. This theme had two subthemes of adequate diversity and inadequate diversity. Participants who identified the diversity of the team as adequate felt that there was adequate representation of the different roles

of individuals within the Academics division of the college. Participant 002 stated “I saw the group was made out of people, which represented various parts of the organization. So, there were people deeply involved in faculty development, but then there were Directors of Academic Affairs, Deans, faculty, adjunct faculty, so I thought that we had brought together good diversity of perspectives.”

Prior Experience

The prior experience theme emerged as a discussion of how the participants’ prior experiences related to their perception of their evaluative thinking skills and development of evaluation skills as a result of participating in the program evaluation. This theme had two subthemes: 1) already practicing, and 2) enhanced skillset. Already practicing pertained to participants who identified, that due to their current role or their background, they were already practicing certain aspects of evaluative thinking or applying a related skill. Examples of these types of statements were from participant 008 who stated, “Given the nature of the positions I’ve held in the past, I felt like a lot of this I have already done and kind of understand.” Or, Participant 003 stated, “I was drawing my answers on other similar situations that I think I asked questions that are thoughtful. So, I was calling on a pattern of what I believe to be true in other work situations.”

Use

The use theme emerged as a discussion of how participants identified ways or desires to increase the use or prevalence of program evaluation. This theme had two subthemes: 1) use in their role within the institution, and 2) desire for increased use as an institutional practice. Use in their current role within the institution pertained to the participant’s desire or expressed ideas to practice program evaluation as a regular routine within their own program or department to collect data and use it for reflection on program or process improvements. Responses which represent this subtheme were ones such as 003 stated, “I definitely, moving forward, can see myself saying, I’m trying to make a determination about this. So how do we want to measure that, and what data points can we pull to start making a decision on it?” Or Participant 008 stated, “My thought has changed from, I don’t have time to try to integrate this, to oh yes, I can incorporate this into smaller components that will eventually result in me being able to evaluate pieces into my current position.”

Conclusion

The purpose of this study was to explore whether or not when limiting the scope of a participatory evaluation using an objectives-oriented evaluation model as an advanced organizer, does the evaluation meet the criteria provided by Daigneault and Jacob (2009) using the Participatory Evaluation Measurement Instrument (PEMI) and Evaluation Involvement Scale (EIS) provided by Toal (2009). The requirement posited by Daigneault and Jacob (2009) for an evaluation to be considered participatory is that the three domains of the PEMI: Extent of Involvement, Diversity of Participants, and Control of the evaluation process, each receive an average rating of at least .25, on a scale of 0.00 – 1.00, when completed by the evaluation team participants. The mean responses from the evaluation participants for each domain were: Extent of Involvement ($M = .6944$), Diversity of Participants ($M = .6667$), and Control of the Evaluation Process ($M = .5278$), indicating that the evaluation was considered to be participatory.

Although there were no statistically significant differences between the mean ratings of the three domains, interview questions were added to the interview protocol to further understand how the participants were interpreting the concepts of involvement, diversity, and control as they relate to the program evaluation and why they selected the rating for each. Participant involvement, team diversity, and control emerged as major themes from the qualitative analysis. Across participants, there was a strong consensus that they felt they were involved in the evaluation process by their engagement in the various evaluation activities. Examples shared of activities included developing data collection instruments, data collection, interpreting results, reviewing work of other participants, collaborated with other participants, and engaged in conversations. No participant responded they were involved in all of the evaluation activities, or none of the evaluation activities. Instead, expressed that they contributed where they saw an opportunity and where they felt comfortable.

The theme of team diversity emerged from responses and descriptions of how the individual participants reported their thinking of diversity of the team as it pertained to the roles represented by the makeup of the team. The roles represented on the team were director of academic affairs, dean, program director, faculty developer, instructional designer, faculty, and practicum site coordinator. Those who reported feeling that there was adequate diversity on the team consistently reported they felt it was representative of the different roles within the Academics division of the college. Those that felt there could have been more diversity of roles on the team identified students and other divisions within the college, such as

Student Affairs, as other roles that would have been appropriate to include. Additional ways in which diversity was discussed were the number of individuals representing each of the roles, full-time and part-time employees, and academic/professional background.

The theme of control emerged from responses discussing the level of control shared between the trained evaluator and the non-evaluative participants. Control was frequently expressed as being shared between the two roles. The various ways in which the non-evaluative participants' control was described were not limited but also not substantial, an ability to access materials, the ability to pick the task they would need to complete and who they collaborated with, and there being a lack of a sense of oversight. The ways in which the evaluator's level of control was described were guiding the process and discussion, keeping the team on task, and helping the team make decisions.

The results of the analysis of the EIS (Toal, 2009) resulted in each evaluation activity item being coded as a little involvement, some involvement, or extensive involvement. The overall mean for the instrument ($M = 2.26$) indicates that the overall evaluation experience for the participants was some involvement. Additionally, the one sample Wilcoxon Signed-Rank Test resulted in six activities having ratings being statistically significant: 1) Identifying evaluation planning team members ($M = 1.56$, $SD = .726$, $p = .009$), 2) Developing data collection instruments ($M = 2.22$, $SD = .833$, $p = .038$), 3) Writing evaluation reports ($M = 1.67$, $SD = 1.00$, $p = .015$), 4) Reviewing evaluation reports for accuracy and/or completeness ($M = 1.56$, $SD = 1.014$, $p = .012$), 5) Presenting evaluation findings ($M = 1.22$, $SD = .667$, $p = .005$), and 6) Developing future project plans based on evaluation results ($M = 1.89$, $SD = 1.167$, $p = .028$). These tasks were represented in the responses within the control and participant involvement themes. Three different participants reported being asked to provide recommendations of participants to add to the evaluation team, which would account for the mean rating between 1 (No involvement) and 2 (Yes, a little). There were several data collection instruments, such as surveys, checklists, and tables, developed as part of the evaluation framework. This provided ample opportunity for participants to be involved in this specific activity. It is also a response which is prevalent in the Participant Involvement theme, which provides insight into why it received a mean rating between 2 (Yes, a little) and 3 (Yes, some).

The final series of activities which were statistically significant were of particular interest due to the relatively high standard deviations compared to the other activities participants responded to on the instrument. It was apparent in the follow-up interview responses, represented in the participant involvement and

OBJECTIVES DRIVEN PARTICIPATORY EVALUATION MODEL

participant control themes, that participants interpreted these activities differently. Some of the participants interpreted these activities as the final reports, presentations, and planning which occurs at the end of the evaluation process, which had not yet occurred at the time they completed the survey. Whereas others interpreted these activities as the preliminary sharing of information which was occurring as the team progressed through the evaluation process, such as documenting findings of specific indicators, sharing data collection instruments with members of the group for feedback, and discussing possible program revisions as sidebars during the program evaluation meetings.

Due to the prevalence of educational programming and the complexity of the different educational programs which exist in educational settings, it makes the likelihood of program evaluations unlikely as institutions are faced with finite resources to invest in their programs. However, institutions are being expected to demonstrate data informed decision making and directors of programs are often expected to demonstrate continuous improvement of their programs (Fitzpatrick, Sanders, & Worthen, 2011). A solution to address this need and support the use of program evaluation within an organization while also being responsible stewards of resources provided to support programs is to engage stakeholders using a participatory evaluation methodology while limiting the scope of the evaluation using program objectives as an advanced organizer.

It is suggested, based on the findings of this study, limiting the scope of a program evaluation using the program objectives helps maintain a focused program evaluation which is approachable for non-evaluative stakeholders. Additionally, implementing the evaluation utilizing a participatory evaluation methodology, has demonstrated that it provides benefits to the participants and the organization by providing training and experience completing a program evaluation with the support of a trained evaluator. While limiting the scope of the evaluation to the program objectives, it may limit the ability of the evaluation team to change the focus or direction of the evaluation; however, it does not prevent the evaluation from being participatory. Additionally, it has the added benefit of promoting the capacity of evaluative thinking within the organization through the participants who had the opportunity to engage in the program evaluation. This has the effect of impacting the participants' thinking, skills, and beliefs about programs and how they are evaluated and revised.

Acknowledgments

This paper is developed in part from the author's unpublished dissertation (Saalman, 2020), and elements throughout are derived from that work.

References

- Alkin, M. (Ed.). (2004). *Evaluation Roots: Tracing Theorists' Views and Influences*. Thousand Oaks, CA: Sage Publications, Inc.
<https://doi.org/10.4135/9781412984157>
- Burke, B. (1998). Evaluating for a change: Reflections on participatory methodology. *New Directions for Evaluation*, 1998(80), 43–56.
<https://doi.org/10.1002/ev.1116>
- Canadian International Development Agency. (2001, March). How to perform evaluations: Participatory evaluations. Retrieved from
<https://www.oecd.org/derec/canada/35135226.pdf>
- Centers for Disease Control and Prevention Program Performance and Evaluation Office. (2018). CDC Program evaluation framework checklist. Retrieved March 9, 2019, from <https://www.cdc.gov/eval/steps/index.htm>
- Cousins, J. B. (2005). Will the real empowerment evaluation please stand up? A critical friend perspective. In D. M. Fetterman & A. Wandersman (Eds.), *Empowerment evaluation principles in practice* (pp. 183–208). NY: Guilford Press.
- Cousins, J. B., & Earl, L. M. (1992). The case for participatory evaluation. *Educational Evaluation and Policy Analysis*, 14(4), 397–418.
<https://doi.org/10.3102/01623737014004397>
- Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. *New Directions for Evaluation*, 1998(80), 5-23. <https://doi.org/10.1002/ev.1114>
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). Thousand Oaks, CA: Sage Publications, Inc.
- Daigneault, P. M., & Jacob, S. (2009). Toward accurate measurement of participation: Rethinking the conceptualization and operationalization of participatory evaluation. *American Journal of Evaluation*, 30(3), 330–348.
<https://doi.org/10.1177/1098214009340580>
- Daigneault, P. M., & Jacob, S. (2013). Unexpected but most welcome: Mixed methods for the validation and revision of the participatory evaluation measurement

OBJECTIVES DRIVEN PARTICIPATORY EVALUATION MODEL

instrument. *Journal of Mixed Methods Research*, 8(1), 6–24.

<https://doi.org/10.1177/1558689813486190>

Fitzpatrick, J. L., Sanders, J. R., & Worthen, B., R. (2011). *Program evaluation: Alternative approaches and practical guidelines* (4th ed.). Boston: Pearson.

Higher Learning Commission. (n.d.). Retrieved February 24, 2019, from <https://www.hlcommission.org/component/directory/?Action=ShowBasic&Itemid=192&instid=1977&lang=en>

Karsh, E., & Fox, A. S. (2014). *The only grant-writing book you'll ever need: Top grant writers and grant givers share their secrets*. NY: Carroll & Graf.

King, J. A., Cousins, B., & Whitmore, E. (2007). Making sense of participatory evaluation: Framing participatory evaluation. *New Directions for Evaluation*, 2007(114), 83–105. <https://doi.org/10.1002/ev.226>

Madaus, G. F., Scriven, M., & Stufflebeam, D. L. (1983). *Evaluation models : Viewpoints on educational and human services evaluation*. Boston, MA: Kluwer Academic Publishers. <https://doi.org/10.1007/978-94-009-6669-7>

Madaus, G. F., & Stufflebeam, D. L. (Eds.). (1989). *Educational evaluation: Classic works of Ralph W. Tyler*. Boston, MA: Kluwer Academic Publishers.

Mertens, D. M. (2018). *Mixed methods design in evaluation*. Thousand Oaks, CA: Sage Publications, Inc.

Moreau, K. A. (2017). Twelve tips for planning and conducting a participatory evaluation. *Medical Teacher*, 39(4), 334–340. <https://doi.org/10.1080/0142159x.2017.1286310>

Neave, H. R., & Worthington, P. L. (1988). *Distribution free tests*. London: Unwin Hyman Inc.

Richlin, L. (2006). *Blueprint for learning: Constructing college courses to facilitate, assess, and document learning*. Sterling, VA: Stylus Publications.

Saalman, D. (2020). *Advancing Objectives-Oriented Evaluation with Participatory Evaluation Methodology: A Mixed Methods Study* [Unpublished Doctoral Dissertation, Wayne State University].

Sorinola, O. O., & Thistlethwaite, J. (2013). A systematic review of faculty development activities in family medicine. *Medical Teacher*, 35(7), e1309–e1318. <https://doi.org/10.3109/0142159X.2013.770132>

Steinmetz, A. (1983). The discrepancy evaluation model. In G. F. Madaus, M. S. Scriven, & D. L. Stufflebeam (Eds.), *Evaluation Models*. Boston, MA: Kluwer Academic Publishers.

DUSTIN R. SAALMAN

Stufflebeam, D. L. (2001). Evaluation models. *New Directions for Evaluation*, 2001(89), 7. <https://doi.org/10.1002/ev.3>

Stufflebeam, D. L., & Coryn, C. L. S. (2014). *Evaluation theory, models, and applications* (2nd ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Stufflebeam, D. L., & Shinkfield, A. J. (2007). *Evaluation theory, models, and applications*. Hoboken, NJ: John Wiley & Sons, Inc.

Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation*, 27(2), 237–246. <https://doi.org/10.1177/1098214005283748>

Toal, S. A. (2009). The validation of the evaluation involvement scale for use in multisite settings. *American Journal of Evaluation*, 30(3), 349–362. <https://doi.org/10.1177/1098214009337031>

Tyler, R. W. (1931). A Generalized technique for conducting achievement tests. *Educational Research Bulletin*, 10(8), 199–208.

Wilcox, R. (2011). *Modern statistics for the social and behavioral sciences: A practical introduction*. Boca Raton, FL: CRC Press. <https://doi.org/10.1201/9781466503236>

Wraga, W. G. (2017). Understanding the Tyler rationale: Basic principles of curriculum and instruction in historical context. *Espacio, Tiempo y Educación*, 4(2), 227–252. <https://doi.org/10.14516/ete.156>

Worthen, B. R., & Sanders, J. R. (Eds.). (1973). *Educational evaluation: theory and practice*. Worthington, OH: C.A. Jones Pub. Co.