Journal of MultiDisciplinary Evaluation
Volume 18, Issue 42, 2022

JMDE
Journal of MultiDisciplinary Evaluation

ISSN 1556-8180
http://www.jmde.com

Empowerment Evaluation of Programs Involving Youth: Evaluators' Perceptions

Sarah Heath
Department of Criminal Justice, University of Winnipeg, Manitoba, Canada

Katherine Moreau
Faculty of Education, University of Ottawa, Ontario, Canada

Background: Participatory and collaborative evaluation approaches, including empowerment evaluation (EE), are useful for evaluating programs involving youth. Empowerment evaluation involves stakeholders in the evaluation process through a set of structured steps. It is primarily concerned with empowering, illuminating, and building program beneficiaries' self-determination.

Purpose: To explore the extent to which evaluators use EE to evaluate programs involving youth, as well as what factor(s) facilitate and hinder their use of EE in these programs.

Setting: The study involved evaluators associated with the collaborative, participatory and empowerment evaluation and youth-focused evaluation topical interest groups (TIGs) of the American Evaluation Association (AEA) who are involved in evaluating programs targeted at youth.

Intervention: Not applicable.

Research Design: The study used a two-phase sequential mixed-methods research design.

Data Collection and Analysis: Phase 1 involved surveys with evaluators, and descriptive statistics (e.g., frequencies and percentages) were calculated for the survey items. Phase 2 included interviews with a sample of evaluators from Phase 1. The interviews were analyzed for common factors.

Findings: In Phase 1, 41 respondents (53.9%) indicated not using EE to evaluate programs involving youth, 30 (39.5%) had used EE, and 5 (6.6%) were unsure. Those who used EE did so to teach youth program stakeholders about evaluation (n = 8, 24.2%), produce more authentic results by engaging youth as experts of their lived experience (n = 7, 21.2%), or produce more useful results for stakeholders to use (n = 6, 18.2%), as well as other less popular reasons. In Phase 2, twelve interviewees raised five factors that facilitate or hinder the use of EE to evaluate programs involving youth. These factors included evaluators' perceptions, evaluators' evaluation experience, evaluators' knowledge and professional training, guidelines from organizations and funders, and stakeholders' availability. Factors that some interviewees viewed as facilitators of EE, others viewed as hindrances.

Keywords: *empowerment evaluation; program evaluation; youth-focused evaluation.*

Introduction

Empowerment Evaluation (EE) is a collaborative approach that involves stakeholders in evaluation through a set of structured steps. The main goal of EE is to build program beneficiaries' self-determination and empower them throughout the evaluation process (Cousins & Whitmore, 1998; Fetterman, 2001). It also aims to teach program beneficiaries how to conduct their own evaluations. In EE, evaluators act as facilitators to assist program stakeholders, which could include program beneficiaries, in designing and implementing program evaluations (Fetterman, 1994). Moreover, they teach stakeholders how to use evaluation processes and results to enact transformative program changes (Fetterman, 1994; Fetterman & Wandersman, 2005; Moreau & Cousins, 2012).

Given the emphasis that EE places on the inclusivity of stakeholders, it appears to be a good fit for evaluating programs that involve youth. According to Langhout and Fernandez (2015), evaluators can use EE to value and respect youth's perspectives and allow them to act as equal partners in evaluation decision-making processes (Beresford, 2000; Fox & Cater, 2011). In such instances, youth have access to evaluators who are designated critical friends, which means they provide evaluation support to stakeholders involved in the EE (Fetterman & Wandersman, 2005; Langhout & Fernandez, 2015; Moreau & Cousins, 2012). In theory, through EE, youth can experience illumination (i.e., revealing or enlightening experiences where new knowledge or new possibilities about roles, structures, and programs become apparent/available) and liberation (i.e., an emancipatory force or freedom from pre-existing roles and constraints, and new conceptualizations of oneself and others) (Fetterman, 1994, 2001). This is because EE can emphasize the development of youth leadership and critical thinking skills, which can aid youth in challenging traditional (and possibly limiting) roles ascribed to them (Flores, 2007; Fox & Cater, 2011). Cumulatively, this evaluative responsibility can build agency for youth as they gain the "ability to create knowledge about the issues and programs that affect their lives" (Zeller-Berkman et al., 2015, p. 25). That is, the information youth obtain by assisting in the design, implementation, and use of evaluation findings can foster feelings of control and self-efficacy, which may help them to cope better in stressful situations.

Scholars have questioned the utility of EE generally and identified issues associated with the use of EE to evaluate programs involving youth. Miller and Campbell (2006) and Cousins (2005) have questioned whether EE is actually different from other collaborative and participatory evaluation approaches. As an approach, it may also pose several challenges when used to evaluate programs involving youth. First, evaluators may need to provide a lot of guidance and training to youth throughout the EE, which may detract from the youth's sense of independence and empowerment. Evaluators may also need to break apart the EE process and develop youth-appropriate instructions for completing each step, which may be time-consuming and challenging (Fox & Cater, 2011). Moreover, while youth need latitude to try out the EE steps, some youth may feel vulnerable trying out the steps or sharing their ideas in front of peers and adults (Fetterman, 1994, 2001; Fetterman & Wandersman, 2005). While EE can encourage youth to embrace transformative social justice and participate in training, facilitation, advocacy, illumination, and liberation exercises, the extensiveness of this involvement may be burdensome or uninteresting to some youth (Fetterman, 2001). In one example detailed by Langhout and Fernandez (2015), youth who were part of the EE learned about challenges present in their school and fixing those issues became an additional hurdle for the youth to overcome. This example illuminates the fact that youth may not always feel empowered by their participation in EE (Miller & Campbell, 2006; Patton, 2005).

Given the above-mentioned potential benefits and challenges of using EE to evaluate programs involving youth, our study aimed to explore the extent to which evaluators use EE to evaluate programs involving youth, as well as what factor(s) facilitate and hinder their use of EE in these programs. The following research questions, as they related to the use of EE, guided this two-phase study:

1. Phase 1: To what extent do evaluators use EE to evaluate programs involving youth?
2. Phase 2: What factor(s) facilitate and hinder the use of EE to evaluate programs involving youth?

Methods

Study Design

The study used a two-phase sequential mixed-methods design. Phase 1 comprised quantitative

survey data from evaluators associated with the American Evaluation Association (AEA)[1] who have an interest in collaborative, participatory, and empowerment evaluation or youth-focused evaluation. The authors collected both quantitative and qualitative (i.e., mixed methods) data sequentially from evaluators in order to describe their use of EE to evaluate programs targeting youth (Creswell, 2014). The authors used the findings from Phase 1 to inform the development of participant-level questions and to identify participants for Phase 2. Phase 2 included qualitative interviews with selected evaluators. Findings from the survey informed the qualitative interview protocol that was used to collect data through semi-structured interviews with a selected group of evaluators (Seidman, 2013). To establish trustworthiness, the authors focused on the concepts of credibility (i.e., confidence) and confirmability (i.e., degree of neutrality) put forth by Lincoln and Guba (1985). The authors triangulated the survey and interview results to build credibility and demonstrate confirmability (Creswell & Plano Clark, 2011; Doyle et al., 2009). Ethics approval was obtained for both Phase 1 and Phase 2 from the University of Ottawa (File number S-05-18-663).

Phase 1: Survey

Sample. This study focused on evaluators from two AEA TIGs, namely (a) Collaborative, Participatory and Empowerment Evaluation and (b) Youth-Focused Evaluation. At the time of the study, there were approximately 746 members in the Collaborative, Participatory and Empowerment Evaluation TIG and approximately 314 members in the Youth-Focused Evaluation TIG (AEA, 2019). Criterion-based sampling (Patton, 2015) allowed the researchers to identify and recruit evaluators who met the criterion of evaluating programs targeting youth. Any evaluator who was using EE or had used it at some point in their career was eligible for the study. Email addresses of members were obtained from the AEA Research Working Group. At the time of the study, the AEA Research Working Group provided the authors with access to the full mailing lists of the TIGs.

Instrument Development. The survey tool draws from our literature review and from the surveys used by other evaluation scholars to ask evaluators to describe their evaluation experiences (for example, Cousins et al., 1995; Sheldon, 2016). The survey comprised 25 questions, including screening questions to check eligibility to participate in the survey, demographic questions, and follow-up-related questions (See Table A1 in the Appendix for the survey specifications). Of the 25, 2 questions were open-ended, while the other 23 were closed-ended. Survey questions included the response option of "I don't know" where relevant to minimize missing responses (Dillman, 2011). In Questions 1 through 3, potential respondents confirmed their eligibility for the study by indicating that they did in fact evaluate programs that involve youth. The survey was built in SurveyMonkey and hosted on a Canadian server to ensure survey data were stored in Canada and therefore subject only to Canadian privacy laws. The survey was piloted with two evaluation colleagues to ensure that it was clearly written and not missing any relevant questions (Lancaster et al., 2004).

Data Collection Procedures. One researcher emailed the information letter and the survey link to the potential respondents. Following Dillman's (2011) tailored design method, the researcher sent respondents the survey link by email three times, including an initial request, a reminder at two weeks, and a final reminder at four weeks from the initial request. Respondents were given the option to skip any questions that they did not wish to answer, which resulted in a different number of respondents for each survey question (See Table A3 in the Appendix for the number of respondents for each question). A total of 108 members completed the survey, constituting a response rate of 13.8%.[2]

Data Analysis. Descriptive statistics (e.g., frequencies and percentages) were calculated for the survey items using Statistical Package for the Social Sciences (SPSS) version 25. Additionally, a

[1] A professional association for evaluators who join voluntarily and pay membership dues to access association services and resources, and attend an annual conference. The AEA includes members from North America as well as other continents.

[2] AEA supplied a list of 746 emails for the Collaborative, Participatory and Empowerment Evaluation TIG and 314 emails for the Youth-Focused Evaluation TIG. When those two lists were combined, 172 emails appeared on both TIG email lists. The duplicate emails were removed, resulting in 888 unique email addresses. After 888 unique emails were sent, 103 were returned as undeliverable or with "out of the office" messages, resulting in 785 functioning email addresses and thus, the potential respondent group for the Phase 1 survey.

content analysis was conducted to analyze the two open-ended survey questions. First, one researcher read the responses for each open-ended question, created a list of reoccurring themes, and re-read the responses and grouped them by the identified themes. The researcher counted the number of responses within each thematic group and calculated counts and percentages for each (Berg, 2008).

Phase 2: Interviews

Sample. The study used criterion-based sampling (Patton, 2015) to identify and recruit evaluators who completed the survey in Phase 1. One researcher contacted the 36 respondents who indicated interest in participating in an interview, provided their email addresses, and indicated whether they did or did not use EE to evaluate programs that involve youth. Of these 36 respondents, 12 (33.3%) replied to our email and agreed to participate in an interview.

Instrument Development. Findings from Phase 1 of the study were used to design the semi-structured interview guides for Phase 2. Separate guides were created for respondents who indicated they used EE to evaluate programs involving youth and for those who indicated they did not. Each guide began with an introductory script and consisted of open-ended questions, including additional probes to allow for expansion. The guide for those who had used EE included 16 questions, while the guide for those who had not used EE included 5 questions (See Table A2 in the Appendix for the interview guide specifications). The guides were piloted with two evaluation colleagues to ensure that they were clearly worded and were not missing any relevant questions or probes (Seidman, 2013).

Data Collection Procedures. One researcher emailed the selected evaluators. The email included the letter of information and consent form. All interviews occurred by Skype or FaceTime and lasted approximately one hour. With the permission of the interviewee, each interview was audio-recorded and transcribed verbatim by one researcher.

Data Analysis. The interview data analysis was informed by Miles and colleagues' (2020) systematic and iterative approach to generating meaning from data and testing and confirming findings. The study used NVivo software (https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home) to track and record the analytic and interpretive process (Varpio et al., 2016). The researcher who coded the data started with "an a priori list of researcher-generated codes" based on the questions asked in the interview guides about context, stakeholder characteristics, organization characteristics, evaluator characteristics, evaluation resources, external factors, and other factors (Miles et al., 2020, p. 69). Subsequently, one researcher read through each transcript three times and then added notations about passages in the transcripts that were related to the codes. Next, both researchers together revised the a priori list; grouped the codes into a smaller number of themes; and transferred the themes, codes, and representative quotations to a data analysis matrix. Initial findings were drafted by reviewing the matrix to develop explanations for the relationships between the codes and generate categories and themes. Both researchers identified exemplar quotes (Miles et al., 2020) to be included in the findings section of this paper, and each interviewee was assigned a pseudonym made up of the letter "E" for evaluator and a number according to the order in which the interview occurred (e.g., E1, E2, etc.).

Findings

Phase 1: Survey

Demographics. Among the respondents who replied to the survey, 67 (62.0%) completed the demographic section. Of the 67 respondents to the demographic section, over half (n = 36, 53.7%) had worked as evaluators for 11 years or more, while 16 (23.9%) had 6 to 10 years of experience and 15 (22.4%) had worked as evaluators for 1 to 5 years. Respondents worked in a variety of areas, including education (25.4%), health (23.9%), multi-disciplinary organizations (20.9%), social research (19.4%), and various other disciplines (10.4%). Of the 108 respondents who participated in the survey, 84 (77.8%) had evaluated programs involving youth, while 24 (22.2%) had not. Of the 84 respondents who noted that they had evaluated programs involving youth, 76 (90.4% of those evaluating programs involving youth, or 70.3% overall) provided information about how many programs involving youth they evaluated throughout their career (depicted in Table 1). Of the 84 respondents who noted that they had evaluated programs involving youth, 75 (89.3%) indicated the number of programs involving youth that they had evaluated over the past year. In the

past year (i.e., from October 2017 to October 2018), respondents had evaluated between zero (n = 10, 13.3%) and 12 programs (n = 2, 2.7%) involving youth. On average, respondents had evaluated nine programs involving youth during the past year.

Table 1. Number of Programs Involving Youth Evaluated by Survey Respondents

Number of programs evaluated involving youth	***n***	%
1 – 3 programs	15	19.7
4 – 6 programs	12	15.8
7 – 9 programs	11	14.5
10 – 12 programs	7	9.2
13 – 15 programs	6	7.9
16 or more programs	25	32.9

Extent of Use of EE for Youth Programs. Of the 84 respondents who noted that they had evaluated programs involving youth, 76 (90.5%) answered the question about their use of EE. Specifically, over half of respondents (n = 41, 53.9%) indicated not having used EE to evaluate programs involving youth, 30 (39.5%) had used EE to evaluate programs involving youth, and 5 (6.6%) were unsure as to whether they had or had not used EE to evaluate programs involving youth.

Respondents who said that they used EE to evaluate programs involving youth or said that they were unsure as to whether they used EE to evaluate programs involving youth were asked why they used EE. Of the 35 respondents who could have answered this open-ended question, 33 provided responses. As shown in Table 2, respondents to this open-ended question used EE to evaluate programs targeting youth for a variety of reasons.

Table 2. Reasons Why Survey Respondents Used EE to Evaluate Programs Targeting Youth

Reasons why EE was used	*n*	%
EE teaches youth program stakeholders about evaluation.	8	24.2
EE produces more authentic results by engaging youth as experts of their lived experience.	7	21.2
EE produces more useful results for stakeholders to implement.	6	18.2
EE aligns with the empowerment and leadership goals of the program evaluated.	4	12.1
EE fully engages program stakeholders in evaluation.	4	12.1
EE helps program stakeholders define the value and utility of their program.	2	6.1
EE helps to earn youth buy-in for evaluation.	1	3.0
EE is in job title.	1	3.0

Note. Percentages do not total to 100 due to rounding.

In terms of what evaluators' EE activities involved and how they involved stakeholders in the EE, respondents reported stakeholders being involved to varying degrees in different evaluation activities, as depicted in Table 3.

Table 3. Evaluators' Perceptions about the Extent to Which Stakeholders Were Involved in Particular EE Activities

Evaluation activities	Level of involvement most frequently reported	% of respondents
Incorporate evidence about their program/project into program/project decision-making	To a great extent	73.9
Identify credible evidence to collect to assess their ability to achieve program/project goals	To a great extent	62.5
Collect their own evidence about their program/project	To a great extent	58.3
Review program/project goals for the future	To a great extent	58.3
Plan program/project goals/benchmarks for the future	To a great extent	56.0
Identify strategies to achieve program/project goals	To a great extent	54.2
Determine the technical knowledge and capacities to collect and analyze evidence of their ability to achieve program/project goals	To a great extent	54.2
Document the current state of their program/project	To a great extent	45.8
Re-assess the current state of their program/project for comparison against a baseline	To a great extent	41.7
Determine strategies to continually collect evidence to assess their ability to achieve program/project goals	A moderate to great extent	45.8
Assess the current state of their program/project to establish a baseline	A moderate to great extent	41.7
Receive training on research methods, including data collection and analysis	To a moderate extent	41.7
Receive training on conducting evaluations	To a moderate extent	37.5
Establish a mission statement for their program/project	To a small extent	37.5

Note. The percentage column's sum is greater than 100 because respondents provided a level of involvement for each activity.

All 84 respondents who noted that they had evaluated programs involving youth were asked about their confidence in using EE to evaluate programs targeting youth. These frequencies are depicted in Table 4.

Table 4. Evaluators' Confidence Using EE

Level of confidence	***n***	%
Very confident	10	13.3
Confident	15	20.0
Somewhat confident	16	21.3
A little confident	21	28.0
Not confident at all	13	17.3

Of these 84 respondents, 74 responded, describing their level of training in EE. The majority (n = 51, 68.9%) of respondents had not been trained in EE, while 21 (28.4%) indicated that they had received training, and 2 (2.7%) were unsure. Of these 84 respondents, 75 indicated their understanding of EE. Close to 30% (n = 22, 29.3%) of respondents rated their understanding of EE at the "advanced beginner" level. The remaining respondents rated their understanding level as "proficient" (n = 17, 22.7%), "novice" (n = 15, 20.0%), "competent" (n = 15, 20.0%), or "expert" (n = 5, 6.7%), and 1 respondent (1.3%) was unsure of their rating. Of the 84 respondents, 74 indicated their experience conducting research on EE. Over half of respondents (n = 47, 63.5%) had not conducted research on EE, 25 (33.8%) said they had, and 2 (2.7%) were unsure if they had or had not conducted research on EE.

The 41 respondents who noted that they did not use EE to evaluate programs involving youth were asked why they did not use EE and which evaluation approaches they used to evaluate programs involving youth. As depicted in Table 5, respondents did not use EE to evaluate programs targeting youth for a variety of reasons. More than half of respondents (n = 23, 56.1%) listed using either practical participatory evaluation or stakeholder-based evaluation to evaluate programs targeting youth, and 18 respondents (43.9%) listed either developmental evaluation or school-based evaluation (see Table 6 for other, less frequent responses).

Table 5. Reasons Provided by Survey Respondents for Not Using EE to Evaluate Programs Targeting Youth

Reasons for not using EE to evaluate programs involving youth	*n*	%
Lack of stakeholder interest	18	43.9
Lack of training and instruction in EE	18	43.9
Lack of clarity around EE	17	41.5
Limited time	16	39.0
Limited funds	14	34.1
Not aligned with evaluation objectives	14	34.1
Not aligned with program context	10	24.4
Never considered it an option	8	19.5
Evaluation client not interested	2	4.9
Lack of influence over evaluation design	2	4.9
Difficulty involving youth in empowerment process	1	2.4
I don't know.	1	2.4
Institutional review board restrictions	1	2.4
The formal application of EE would have little benefit beyond what is currently being used.	1	2.4
Youth are young, so unclear how it would work.	1	2.4

Note. The percentage column's sum is greater than 100 because respondents could select more than one response.

Table 6. Approaches Used by Survey Respondents to Evaluate Programs Targeting Youth

Approaches used to evaluate programs involving youth	*n*	%
Practical participatory evaluation	23	56.1
Stakeholder-based evaluation	23	56.1
Developmental evaluation	18	43.9
School-based evaluation	18	43.9
Theory-driven evaluation	4	9.8
Transformative participatory evaluation	4	9.8
Democratic evaluation	2	4.9
I don't know.	2	4.9
Utilization-focused evaluation	2	4.9
Active evaluation	1	2.4
General participatory evaluation (not T-PE or P-PE)	1	2.4
Realist evaluation	1	2.4
Whatever suits client	1	2.4

Note. The percentage column's sum is greater than 100 because respondents could select more than one response.

Phase 2: Interviews

Demographics. Of the 84 respondents who completed the Phase 1 survey and indicated that they evaluate programs involving youth, 36 (42.9%) agreed to be contacted about participating in an interview. Among the 36 individuals who agreed to be contacted to participate in an interview, 12 consented to be interviewed. Half of the interviewees (n = 6) indicated they had used EE to evaluate programs involving youth, while the other half (n = 6) had not. At the time of the interview, nine (75.0%) interviewees resided in the United States, two (16.7%) resided in Canada, and one (8.3%) resided in a European country. All interviewees were employed as full-time evaluators. Eight interviewees (66.7%) held senior evaluation positions with 10 to 22 years of evaluation experience, two interviewees held mid-level positions with 4 to 5 years of experience, and two interviewees classified themselves as junior evaluators with 1 to 2 years of evaluation experience. Half of the interviewees (n = 6) worked as internal evaluators at non-governmental organizations, while the other half (n = 6) were self-employed as evaluation consultants.

Factor(s) Facilitating or Hindering Use of EE for Youth Programs. The sections that follow describe the factors that the interviewees believed facilitated or hindered the use of EE for evaluating programs involving youth: (a) evaluator's perceptions, (b) type of evaluation experience, (c) evaluator's knowledge and professional training, (d) guidelines from organizations and funders, and (e) stakeholders' availability. Factors that some interviewees viewed as facilitators, others viewed as hindrances.

Evaluator's Perceptions. Interviewees viewed their perceptions as evaluators as either facilitating or hindering their use of EE to evaluate programs involving youth. Interviewees who indicated they use EE for programs involving youth described how their positive perceptions toward EE facilitated their use. These interviewees believe that EE allows for ongoing feedback from stakeholders in evaluation processes and provides organizations with important information that they can use to improve their programs. E1 noted that "EE provides immediate feedback that [organizations] can use for daily decision making," which E2 described as "pretty important to understanding the important changes [organizations] should make to programming." Interviewees also commented on how EE can build evaluation capacity and empower stakeholders. E5 remarked that evaluators can "build stakeholders' capacity to do and use evaluation through the EE process," and E1 noted that the enhanced capacity enables stakeholders to "go on and continue evaluating when [the evaluation] finishes." Through the EE process, these interviewees believe that they can "make sure [stakeholders'] orientation to evaluation is productive, enthusiastic, and excit[ing]" (E10). One interviewee explained that "EE empowers people" by emphasizing stakeholder feedback:

> EE empowers people because it is related to conversations about vulnerable populations, inclusiveness, and disparities. Like there are these groups that are really impacted by these problems and EE comes at that problem by asking, 'how are we going to understand them better,' by asking them. (E11)

These interviewees reflected that they felt connected to EE through their own "essential beliefs and values about people and society" (E6), and that they felt an "ethical and moral responsibility to conduct EE" (E2). They discussed how EE meets their "personal desire" to contribute to society: "I use EE because I have a personal desire to leave the client with tools for when the evaluation ends… to be able to be reflective for their project themselves" (E1). As another interviewee noted:

> When you're an empowerment evaluator, you're really out there for the social justice piece, and I made my way to EE because the more I refined what I want out of my life, the more I had a personal desire to do EE. (E4)

Moreover, these interviewees perceive EE as an opportunity to cocreate and collaboratively conduct program evaluation alongside stakeholders. They believe that "the more involved [evaluators] can get stakeholders in the process, the better [the evaluation] product and processes will be" (E6). These interviewees disclosed that they view "stakeholders as capable to evaluate and judge their own performance" (E7) and that they involve stakeholders in evaluation because of their unique expertise. As one interviewee said, "I involve stakeholders because they are the ones that know what they need; I'm just making stuff up in my office" (E10).

However, other interviewees who indicated that they do not use EE flagged that their prior preconceived negative perceptions toward EE hinder their use of it to evaluate programs involving youth. They discussed perceiving EE as not useful to the majority of clients, who want to prove that their anticipated outcomes are achieved. These interviewees suggested that most of the clients they interact with have "an exclusive focus on measuring outcomes" (E9) or collecting data on "the bottom line" (E12), and that "EE can't really tell you if changes actually took place" (E3). Interviewees also said that they perceive EE to be a biased form of evaluation (i.e., overly positive or negative) due to the involvement of stakeholders in the evaluation and the potential for that involvement to sway the evaluation, thus hindering evaluators from using it. As one interviewee explained, "most organizations want the evaluation and evaluator to be at arm's length away from the organization, so that the evaluation is seen as unbiased and external and separate from the people and work of the organization" (E9). In addition, they noted that they think busy stakeholders "don't want to be involved in an EE" (E8). These interviewees said that they perceive that the stakeholders are not interested in deep involvement in the evaluation, and instead prefer to hire an external/third-party evaluator to conduct the evaluation on their own. One interviewee described the desire of organizations to hire "a standard external evaluator that's very hands-off to outsource the whole evaluation, because the organization has a lot on their plate, so they want someone who's going to do the job and get it off their list" (E3). Overall, interviewees appeared to be influenced to use EE or not use EE due to their positive or negative perceptions of EE.

Type of Evaluation Experience. The interviewees reflected a similar pattern based on interviewees' type of evaluation experience. Interviewees viewed their own and their colleagues' previous evaluation experience as a factor in deciding whether or not to use EE. Interviewees commented that their past work involved supporting stakeholders in the design and delivery of evaluations, which allowed them to develop facilitation and interpersonal skills (e.g., communication, negotiation, judgment, tact, etc.). These interviewees highlighted how their previous evaluation experience facilitates their use of EE. They said that strong interpersonal and facilitation skills are essential for conducting EE. These interviewees explained that interpersonal skills are important for EE to "get at the populations that are being served, and getting people involved in the evaluation" (E10). They explained that facilitation skills are important to allow the EE evaluator "to be comfortable enough to be a critical friend; meaning they are reflective and make suggestions based on reflections along the way" (E8). They also described an EE evaluator as "someone who has been a 'jack of all trades' and knows how to be whatever the group needs [them] to be" (E1).

Conversely, other interviewees stressed how their previous experiences as evaluators hindered their abilities to use EE. Interviewees said that their past evaluation experience involved working independently in the collection, analysis, and dissemination of evaluation findings, rather than working collaboratively. As one interviewee noted, "Not everyone is qualified to be an empowerment evaluator. Being an empowerment evaluator means moving from a third-party external evaluator who's in control to an ally who is able to support others in their fight where they need

support" (E4). These interviewees further stated that they have seen their colleagues involve stakeholders as data sources in evaluation, but that these stakeholders were not involved to the extent that EE requires. Another interviewee explained that limited contact with stakeholders hinders EE: "I find EE doesn't work well if you're not planning on having many touchpoints with your stakeholders… so communication skills and talking to people and being able to say 'What do you think?' really matters" (E1). In summary, interviewees' past experience conducting EE, or the absence of that experience, appeared to have an impact on their use of EE.

Evaluator's Knowledge and Professional Training. Similarly, interviewees viewed their own and their colleagues' knowledge and professional training as a factor in deciding whether or not to use EE. Interviewees who use EE to evaluate programs involving youth disclosed that they learned about EE from university-instructed courses, by "work[ing] with another evaluator who was familiar with EE" (E2), or by attending an AEA-sponsored EE talk. However, interviewees who do not use EE viewed their lack of knowledge and professional training on EE as a hindrance to using it to evaluate programs involving youth. These interviewees shared the perception that "most people don't know about EE" (E12) and explained that they, as evaluators, would be hesitant to use an unfamiliar evaluation approach. One interviewee remarked, "If an evaluator doesn't know about EE, why would they ever use it in an evaluation? You'd just look incompetent" (E9). Interviewees noted, for example, knowing very little about what EE is, how it differs from other evaluation approaches, what's so empowering about it, when it is appropriate and not appropriate to use, the skills they would need to carry out such an evaluation, and how to do so. Several interviewees also observed that EE is a more specialized approach, so knowledge of how it works is limited to those who have specifically sought out information and training on it. In speaking about their lack of knowledge, one interviewee attributed this insight to an absence of exposure "to anything on EE" (E3). These interviewees noted that "EE requires specialized knowledge and training that not every evaluator would have, meaning that [EE is] not something you can just throw on someone to do" (E8). Interviewees argued that they would only use EE if they felt they had enough information on how to conduct an EE from a combination of training and first-hand experience.

Guidelines from Organizations and Funders. Interviewees viewed guidelines from organizations and funders as a factor in deciding whether or not to use EE. Interviewees noted that organizations and funders facilitate their use of EE by making it a requirement or by leaving the choice of the evaluation approach to the evaluator. These interviewees listed examples of organizations or funders promoting the use of EE "as a type of evaluation that is philosophically congruent with inclusiveness and equality-building" (E4). They cited using EE in circumstances where "the funds dedicated to the evaluation explicitly require the use of EE" (E2), or when the evaluator contract was posted "under the title of empowerment evaluator" (E7). They expressed that the guidelines requesting the use of EE or requests from organizations and funders to focus on stakeholders in the evaluation led them to select an EE approach. These interviewees described using EE in response to organizational and funder demands that the evaluation "involve and empower stakeholders" (E10), "build stakeholder capacity" (E6), or "respect that underserved populations have a voice and are experts in their own lives" (E5). One interviewee explained how organization requests to "listen to and involve" stakeholders led to their use of EE for the first-time: "I first used EE because the organizational contact we worked with was one of those superintendents that was consistently looking for ways to listen to and involve the students and families they served" (E1). These interviewees further commented on how a lack of any guidelines had also led them to use EE for the evaluation of programs involving youth, even if it was not mandatory.

Conversely, interviewees mentioned that organizations and funders hindered their use of EE to evaluate programs involving youth by restricting them to specific evaluation approaches other than EE. As one interviewee explained:

> I would have liked to use EE in my last evaluation, but [the organization is] funded by the federal government and they had a mandate for specific data elements that we had to collect, so that structure [for the evaluation] was set even before [the evaluation] got to us. (E2)

Interviewees commented that they treated organization and funder requirements as mandatory and did not stray from the approaches requested by these groups. One interviewee expressed concern about departing from the organization or funder's terms of reference (i.e., guidelines) for the evaluation, "[the evaluation approach is] whatever the client wants, and sometimes that leaves me with little to no choice in

how to structure the evaluation, but those terms must be followed, or I've violated my agreement" (E12). These interviewees also stated that they were excluded from using EE if the client desired that the evaluation focus on demonstrating outcomes. As one interviewee stated, "In all of those [evaluations], a theory-based approach was used because the client wanted to demonstrate to the funder accountability and if they didn't get to the outcomes, why did that happen" (E3). Evaluators reported feeling it was necessary to follow these requirements set forth by funders and organizations.

Stakeholders' Availability. Interviewees viewed stakeholders and time as either facilitators or hindrances to using EE to evaluate programs involving youth. Interviewees who use EE discussed stakeholders and their time allotted for the evaluation as a factor that facilitates the use of EE to evaluate programs involving youth. They focused on the amount of time available to recruit, coordinate, and involve stakeholders for the EE and the willingness of stakeholders to assume such responsibilities. As one interviewee explained, "Think of it like you're doing a lot of back-and-forth between people, so there's a lot of coordination that's needed; and so, you have to consider the fact that everything is going to take longer" (E2). Interviewees who use EE explained how they would ask project managers and organizational executive directors to coordinate the EE sessions by identifying the appropriate stakeholders who needed to attend, preparing those stakeholders, and then handling the logistics of getting groups of stakeholders to the various EE sessions. They suggested that this coordination function facilitated the use of EE to evaluate programs involving youth by getting stakeholders to the EE and encouraging their participation in the EE. These interviewees viewed EE "as an investment that requires time and support to do it right" (E7). They commented that they would not use EE if the evaluation had "limited support staff and a tight timeline" (E10). These interviewees believe EE requires "more dedicated time and staff because you're not only doing the evaluation, but you're trying to build in capacity to do it in a way that requires teaching people things" (E4).

However, interviewees who do not use EE to evaluate programs involving youth believed that the lack of dedicated stakeholders and time hindered their use of EE. These interviewees detailed their lack of access to stakeholders who would carry out coordination tasks, or their lack of time to use EE to evaluate programs involving youth. They discussed the tight timelines of their evaluation contracts' hindering their use of EE. These interviewees also commented that in the organizations that they have worked there is nobody who can assume responsibility for coordinating the EE. They explained that busy project managers and executive directors do not have time to coordinate participants for an EE, and thus it would be unrealistic to assign such stakeholders the task. As one interviewee observed, "It's a tough job in the short run selling [EE] to overworked teachers and administrators who are focused on providing mainly administrative data to funders to prove how they're spending their time" (E9). These interviewees commented on how the organizations and funders they interact with are focused on "just getting finished a deliverable like an evaluation plan or a report, and so there's not capacity or time for an EE to be done" (E12).

Discussion

This study explored the use of EE to evaluate programs involving youth and the facilitators and hindrances of the use of EE to evaluate programs involving youth. The majority of Phase 1 respondents indicated that they did not use EE to evaluate programs involving youth. This lack of use aligns with, and may be driven in part by, the lack of published material on the use of EE with youth. Fetterman (2001) first introduced EE approximately 20 years ago, yet few researchers have published works on the use of EE with youth since its introduction (see Langhout and Fernandez, 2015, for one example of such a published work). This lack of peer-reviewed work does not bode well for the use of EE as a way of evaluating programs involving youth. For instance, evaluators may look to previously published examples and articles to understand whether EE is appropriate and feasible before using it to evaluate their programs (Mark, 2008; Smith, 1993). In contrast, there are numerous publications on the use of EE to involve adult stakeholders in program evaluation (see for example the edited collections by Fetterman & Wandersman, 2005; Fetterman et al., 2018). Without relevant published material, evaluators may have difficulty acquiring the knowledge and confidence to conduct EE to evaluate programs involving youth.

Indeed, most evaluators in this study perceived their level of knowledge on EE as only "advanced beginner." They also noted only "a little confidence" in their EE abilities. Moreover, evaluators in Phase 2 commented on their lack of knowledge about EE as a hindrance to using it to

evaluate programs involving youth. They noted that they would not use an evaluation approach that they did not know, suggesting that their knowledge directly related to their confidence in using an approach. Those who did feel knowledgeable linked this feeling to the formal training they had received in EE, suggesting that knowledge of and confidence in using EE could be improved through increased instruction in EE by professional associations and educational institutions that train evaluators. However, the work of some evaluation scholars also suggests that a lack of knowledge about EE may not be solved through additional training, because understanding what is EE is, is complex, and because the concept of EE lacks conceptual clarity (Cousins, 2005; Miller & Campbell, 2006).

According to the EE literature and those who have studied it, there appear to be different ways of defining and practicing EE. Fetterman and Wandersman (2005) argue that EE has 10 central principles that distinguish it from other forms of collaborative and participatory evaluation in theory and practice. Table 7 shows the principles identified by Fetterman and Wandersman (2005) and a description of each principle.

Table 7. Empowerment Evaluation Principles

Principle	Description
Improvement	Build on substantive and relevant issues
Community ownership	Values and facilitates community control
Inclusion	All contributions are welcome
Democratic participation	Open and fair decision-making
Social justice	Evaluation is useful to address social inequalities
Community knowledge	Respects and values community knowledge
Evidence-based strategies	Respects and uses knowledge base of scholars
Capacity building	Enhances stakeholders' ability to prepare evaluation and use it to improve programming
Organizational learning	Evidence of use of evaluation to build new practices to inform decision-making, implement program practices, and help organizations learn from experience
Accountability	Outcomes function within existing policies, standards, measures of accountability

However, Miller and Campbell (2006) and Cousins (2005) note that these 10 principles are indicative of collaborative and participatory evaluation in general and are not specific to EE. Additionally, the challenge with using theoretical principles (as opposed to concrete actions) to distinguish evaluation approaches is that principles are open to interpretation and application. Indeed, in their examination of published accounts of EE, Miller and Campbell (2006) found wide variation among practitioners' adherence to the EE principles, which they attribute to conceptual ambiguity. In contrast, Fetterman et al. (2018) perceive that the flexibility of these principles is advantageous as it allows evaluators to adapt the latter to their local context.

Fetterman (2001) also argues that EE differs from other collaborative and participatory approaches in that it necessitates that evaluators engage stakeholders in a series of three crucial steps: (1) developing a mission, (2) taking stock, and (3) planning for the future. Yet, evaluation scholars have also published alternative models for conducting EE (see for example the Getting To Outcomes 10-step model by Wandersman et al., 2000). Evaluators in our study were never directly asked to define EE; however, in Phase 1 evaluators told us about the activities that they used to involve stakeholders in their empowerment evaluations. Evaluators noted they involved stakeholders in taking stock (Step 2 of 3) to a moderate to great extent and in planning for the future (Step 3 of 3) to a great extent, but stakeholders were only involved in developing a mission (Step 1 of 3) to a small extent. Fetterman argues that EE can be distinguished from other

collaborative and participatory evaluation approaches by use of these three steps, but it appears that some steps are used more than others and that they may not, in practice, always be used as a three-step process. In total, evaluators who responded to the survey used 13 different evaluation activities to involve stakeholders in EE. While some of the activities directly relate to the three steps, others are ways of enacting the EE principles and may reflect collaborative and participatory evaluation approaches in general. Therefore, our study would appear to provide further support for the work of Miller and Campbell (2006) about the wide variation in how EE is practiced. The study also highlights ways of practicing the EE principles.

Yet, some of the principles that Fetterman and Wandersman (2005) suggest are markers of EE also appeared to overlap with the reasons provided by evaluators for using EE. In both Phase 1 and Phase 2, evaluators noted using EE to teach youth program stakeholders about evaluation. This reason directly aligns with the principle of capacity building, described by EE scholars as enhancing stakeholders' ability to prepare evaluations and use evaluations to improve programming. Capacity building appears in the general literature on the use of participatory and collaborative evaluation approaches to evaluate programs involving youth (Checkoway & Richards-Schuster, 2003; Flores, 2007; Fox & Cater, 2011; Moreau & Cousins, 2012; Upshur & Barreto-Cortez, 1995; Samuelson et al., 2013; Zeller-Berkman et al., 2015). In general, participatory and collaborative approaches are popular with evaluators because, with assistance from a trained evaluator, stakeholders are involved in the design, implementation, and dissemination of evaluations (Moreau & Cousins, 2012; Upshur & Barreto-Cortez, 1995). Indeed, "a central premise of EE is that programs are more likely to achieve desired outcomes if key stakeholders have the capacity to conduct and use their own evaluations" (Wandersman et al., 2015, p. 646). According to the EE literature, EE adds to the capacity-building process by providing stakeholders access to an evaluator, who is a critical friend with education and experience in evaluation to support the stakeholders (Fetterman & Wandersman, 2005; Langhout & Fernandez, 2015; Moreau & Cousins, 2012). Evaluators in Phase 2 of the study spoke about EE's effectiveness for building evaluation capacity as an important factor in creating their positive perception of EE, which encouraged their use of EE.

Capacity building was discussed as a way to create positive orientations to evaluation and increase the likelihood that evaluation is used in the future, and also as a way to empower stakeholders. In this way, capacity building is also linked to the second popular reason for using EE, as reported by the survey respondents of this study: to produce more authentic results by engaging youth as experts of their lived experience. Unlike capacity building, this reason doesn't appear on the EE principles list as an explicit principle on its own. Instead, this idea seems to connect and relate to the EE principles of (a) community ownership (i.e., valuing and facilitating community control); (b) inclusion (i.e., welcoming all contributions); (c) democratic participation (i.e., open and fair decision-making); (d) social justice (i.e., evaluation as useful to address social inequalities in society); and (e) community knowledge (i.e., respect and valuing of community knowledge).

Similarly, in Phase 2, evaluators spoke about the ability to build the evaluation capacity of program stakeholders, including those in vulnerable positions or who are often left out of conversations that impact their own lives. Empowerment evaluation encourages these groups to participate in the evaluation such that they are in control of important evaluation decisions. This idea is apparent in the collaborative and participatory evaluation literature more widely, which suggests that the consequence of such involvement is that evaluation data are grounded in stakeholders' perspectives (Patton, 1997). This grounding encourages the collection of unique knowledge about programs, including necessary program improvements and key areas of inquiry (Moreau & Cousins, 2015; Upshur & Barreto-Cortez, 1995). Moreover, when such approaches are used to evaluate programs involving youth, youth can learn about program evaluation and offer unique insights about the program from their particular point of view as youth program beneficiaries (Checkoway & Richards-Schuster, 2003; Zeller-Berkman et al., 2015). So, collaborative and participatory evaluation approaches are particularly useful when there is reason to believe that a stakeholder group would hold perspectives that may differ from the organization's, funder's, program staff's, and evaluator's. As an illustration, Zeller-Berkman et al. (2015) discovered that "collaboration with youth on an evaluation survey brought up issues that would have never come to mind for the adult staff members" (p. 28). Moreover, such approaches are promising for helping service providers identify findings that are meaningful from the point of view of youth, and for making programs more meaningful for disadvantaged youth (e.g., youth with mental

health challenges: Bulanda et al., 2013; Chen et al., 2010; Dold & Chapman, 2012).

In addition to the benefit of this involvement for evaluations, the EE literature on adult involvement and the collaborative and participatory literature on youth involvement speak about the effects of this involvement on a program's intended beneficiaries. This literature suggests that those involved in EE experience illumination and liberation and develop leadership and critical thinking skills (Fetterman, 1994, 2001; Flores, 2007; Fox & Cater, 2011). Additionally, the information that youth obtain by assisting in the design, implementation, and use of evaluation findings could foster feelings of control and self-efficacy (Zeller-Berkman et al., 2015). Youth could then take on non-traditional roles (e.g., as co-evaluators) in their program communities. These sentiments were expressed by evaluators in Phase 2 who spoke about their work as empowerment evaluators as contributing to important beliefs and values about people and society and social justice causes by viewing stakeholders as not only capable, but as experts. However, whether youth experienced transformation as a result of the EE was not addressed in this study. Evaluators in this study did mention that the potential for such experiences is dependent on the evaluator's ability to access and recruit youth, the organization's and funder's support for involving stakeholders in the evaluation to a significant degree, and the stakeholder time and resources allotted for the evaluation. So, while the EE literature may suggest such experiences and transformations are possible, this study did not explore whether such outcomes occur when EE is used to evaluate programs involving youth. Yet, the study did examine the conditions (i.e., facilitators and hindrances) that make the use of EE more or less likely.

The influential role played by organizations and funders was explained by evaluators in detail during the Phase 2 interviews. Consistent with the collaborative and participatory literature, evaluators said that the intentions and desires of organizations and funders drive their use of particular evaluation approaches, which is one of the main reasons why the use of EE is not mainstream in this area (Barrington, 1999; Checkoway & Richards-Schuster, 2003; Fox & Cater, 2011; Miller & Campbell, 2006; Samuelson et al., 2013; Zeller-Berkman et al., 2015). In order to facilitate such use, Fetterman et al. (2018) suggest the need for evaluators to provide clarity and guidance to "community members and funders, ensuring that an appropriate match is made between the evaluation approach and the specific needs and resources of the community" (p. v). Furthermore, Fetterman et al. (2018) argue that evaluation scholars have a role to play by advancing the "conceptual clarity and methodological specificity" (p. vi) of EE through practice, documentation, and dissemination with the help of professional bodies and scholarly communities in the field of evaluation.

With these considerations, funding agencies could not only require the use of EE to evaluate programs, but also require and provide funding and time to support the EE. Similarly, evaluator competencies, like those required for the Canadian Evaluation Society Credentialed Evaluator designation, may introduce a competency related to knowledge of and experience with using collaborative and participatory evaluation approaches to evaluations of youth programs. Such a competency would be useful for organizations looking for evaluators and for evaluators wishing to develop their evaluation skillsets for future work evaluating programs involving youth.

According to both the Phase 1 and Phase 2 findings, there are also other barriers to using EE to evaluate programs involving youth. These barriers are similar to those that exist for collaborative evaluation approaches in general and are also similar to those for involving adults in EE. Evaluators in the current study who did not use EE attributed their reluctance to negative perceptions of EE. In the Phase 1 survey, evaluators who did not use EE noted multiple reasons, the most popular being lack of stakeholder interest, lack of training and instruction in EE (as discussed above), and lack of clarity around EE (also discussed above), as well as limited time, limited funds, and a lack of alignment with the objectives of the evaluation. Instead, these evaluators looked to other types of collaborative or participatory evaluation approaches. In Phase 2, evaluators who did not use EE specified that EE is ineffective for clients who want to prove that outcomes are achieved and want an unbiased and detached evaluation (an argument that was also made by Miller and Campbell, 2006). Evaluators added that their own limited contact with stakeholders during the evaluation process, along with organizations' and/or funders' lack of interest to involve stakeholders, held them back from using EE to evaluate programs involving youth.

Limitations and Future Directions for Research

This study explored how evaluators use EE to evaluate programs involving youth, as well as the facilitators and hindrances of EE use. Future research should examine the extent and nature of youth involvement in empowerment evaluations of youth programs, as well as factors that hinder or facilitate their involvement. Future research should also focus on youth's perspectives about being involved in EE.

The authors did not ask respondents how they define EE, and how evaluators define EE can influence the extent to which they use it, as well as the factors that facilitate and hinder their use of it. Future studies should explore how evaluators define EE and examine whether their definitions of it are consistent with the theoretical literature on EE.

Additionally, the survey response rate of 13.2% can be viewed as a limitation of this study. While it is plausible that additional responses would have yielded different results, the response rate is consistent with other surveys targeted at evaluators and run through the AEA TIG listings. For example, one recent surveyor of AEA evaluators indicated that "despite research suggesting that AEA members consider research on evaluation as important, response rates for research on evaluation studies are often only between 10–30%" (Wanzer & Wisner, 2020, para. 1). Similarly, the TIG distribution lists were provided by AEA, and so their accuracy to represent all evaluators who evaluate programs involving youth may be questionable and may have resulted in coverage errors, such as failing to include some portion of the targeted population (Lee, 2008). Moreover, examinations into survey response rates have suggested that online surveys have lower response rates (Kaplowitz et al., 2004) and that acceptable response rates differ across studies and disciplines (Carley-Baxter et al., 2009). However, future research in this area could work on increasing this response rate.

Another issue related to the survey is the potential that evaluators self-excluded from the survey due to the title (i.e., self-selection bias; Cook & Campbell, 1979). Upon receiving the survey, some evaluators (including some with experience in collaborative evaluation) told us they had decided that the survey was not relevant for them due to the term "empowerment evaluation" in the title. These individuals would have contributed to a higher response rate if they had completed the survey and allowed the survey to eliminate them, rather than not taking part at all. Likewise, only 62% of the respondents completed the demographics section of the survey. This lack of demographic information is unfortunate, as without it we could not explore fully whether evaluators' demographic characteristics influenced their use of EE.

Phase 2 relied on evaluators' reports of their perceptions. Respondents may have provided socially desirable answers, and some interviewees may have edited their comments in order to reflect themselves and their interests in a more favourable light. There is also the possibility that the interviewees who agreed to be interviewed are a particular group of evaluators and do not represent evaluators in general, resulting in possible selection bias (Cook & Campbell, 1979). That is, there are suspected differences between those who participated and those who did not. It is likely that evaluators who participated in this phase were more vocal, active in their professions as evaluators, and interested in the topic areas of youth evaluation and/or EE. Thus, these particular evaluators may have expressed different perceptions than those who did not participate. Additionally, since only one researcher coded the data, the study does not report on inter-rater agreement. This use of one researcher may have impacted the trustworthiness of the analysis.

Recalling the words of Smith (1993), research on evaluation has the ability to alter evaluation practice. This study has illuminated areas where further research is needed to advance evaluation practice. As mentioned, the survey was limited to AEA TIG members; future research could reach outside this single professional community and survey evaluators outside the AEA or CES membership lists. Such studies could also be done with a more general study title that may entice evaluators to participate in data collection activities.

Overall, this study sought to understand the use of EE to evaluate programs involving youth. Through the use of mixed methods, including surveys and interviews, it examined the perceptions of evaluators on the topic. The findings show that while the use of EE to evaluate programs involving youth may be limited, there are factors that both facilitate and hinder the use of EE. This work spurred us to collect further empirical research on evaluation to confirm, contradict, and expand our notions on the use of EE to evaluate programs involving youth. We hope this work will aid evaluators in considering the appropriateness of EE for their youth-focused evaluations.

Author Note

Conflicts of Interest. There are no conflicts of interest for any of the authors.

Funding. The Ontario Graduate Scholarship and the Centre for Research on Educational and Community Service funded this study.

Corresponding Author. Dr. Sarah Heath, Department of Criminal Justice, University of Winnipeg, 515 Portage Avenue, Winnipeg, Manitoba, Canada, R3B 2E9; s.heath@uwinnipeg.ca .

References

American Evaluation Association. (2019). *Topical interest groups*. https://comm.eval.org/communities/allcommunities

Barrington, G. V. (1999). Empowerment goes large scale: The Canada prenatal nutrition experience. *The Canadian Journal of Program Evaluation*, *14*, 179–192.

Beresford, P. (2000). Service users' knowledges and social work theory: Conflict or collaboration? *British Journal of Social Work*, *30*(4), 489–503. https://doi.org/10.1093/bjsw/30.4.489

Berg, B. (2008). *Qualitative research methods for the social sciences* (7th ed.). Pearson Education.

Bulanda, J. J., Szarzynski, K., Siler, D., & McCrea, K. T. (2013). "Keeping it real": An evaluation audit of five years of youth-led program evaluation. *Smith College Studies in Social Work*, *83*(2–3), 279–302. https://doi.org/10.1080/00377317.2013.802936

Checkoway, B., & Richards-Schuster, K. (2003). Young people as competent citizens. *Community Development Journal*, *38*(4), 298–309. https://doi.org/10.1093/cdj/38.4.298

Chen, P., Weiss, F. L., & Nicholson, H. J. (2010). Girls study Girls Inc.: Engaging girls in evaluation through participatory action research. *American Journal of Community Psychology*, *46*(1–2), 228–237. https://doi.org/10.1007/s10464-010-9328-7

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design & analysis issues for field settings*. Rand McNally College Publishing Company.

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). Guilford Press.

Cousins, J. B., Donohue, J. J., & Bloom, G. A. (1995). Collaborative evaluation in North America: Evaluators' self-reported opinions, practices and consequences. *Evaluation Practice, 17*(3), 207–226. https://doi.org/10.1177/109821409601700302

Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. *New Directions for Evaluation*, *1998*(80), 5–24. https://doi.org/10.1002/ev.1114

Carley-Baxter, L.R., C.A. Hill, D.J. Roe, S.E. Twiddy, R.K. Baxter & Ruppenkamp, J. (2009). Does response rate matter? Journal editors use of survey quality measures in manuscript publication decisions. *Survey Practice. 2* (7), 1-7.

Creswell, J. W. (2014). *A concise introduction to mixed methods research*. SAGE Publications, Inc.

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Sage Publications.

Dillman, D. A. (2011). *Mail and internet surveys: The tailored design method* (2nd ed.). John Wiley & Sons.

Dold, C. J., & Chapman, R. A. (2012). Hearing a voice: Results of a participatory action research study. *Journal of Child and Family Studies*, *21*(3), 512–519. https://doi.org/10.1007/s10826-011-9505-9

Doyle, L., Brady, A. M., & Byrne, G. (2009). An overview of mixed methods research. *Journal of Research in Nursing, 14*(2), 175–185. https://doi.org/10.1177/1744987108093962

Fetterman, D. M. (1994). Empowerment evaluation. *Evaluation Practice*, *15*(1), 1–15. https://doi.org/10.1177/109821409401500101

Fetterman, D. M. (2001). *Foundations of empowerment evaluation*. Sage Publications.

Fetterman, D. M., Rodríguez-Campos, L., & Zukoski, A. (2018). *Collaborative, participatory, and empowerment evaluation*. Guilford Press.

Fetterman, D. M., & Wandersman, A. (Eds.). (2005). *Empowerment evaluation principles in practice*. Guilford Press.

Flores, K. S. (2007). *Youth participatory evaluation: Strategies for engaging young people*. Jossey-Bass.

Fox, J., & Cater, M. (2011). Participatory evaluation: Factors to consider when involving youth. *Journal of Extension*, *49*(2).

Kaplowitz, M. D., Hadlock, T. D. & Levine, R. (2004). A comparison of web and mail survey response rates. *Public Opinion Quarterly, 68*(1), 94-101.

Lancaster, G. A., Dodd, S., & Williamson, P. R. (2004). Design and analysis of pilot studies: Recommendations for good practice. *Journal of Evaluation in Clinical Practice, 10*(2), 307–312. https://doi.org/10.1111/j..2002.384.doc.x

Langhout, R. D., & Fernandez, J. S. (2015). Empowerment evaluation conducted by fourth- and fifth-grade students. In D. M. Fetterman, S. J. Kaftarian, & A. Wandersman (Eds.), *Empowerment evaluation: Knowledge and tools for self-assessment, evaluation capacity building, and accountability* (2nd ed., pp. 193–232). SAGE Publications, Inc. https://dx.doi.org/10.4135/9781483387079

Lee, S. (2008). Noncoverage. In P. J. Lavrakas (Ed.), *Encyclopedia of survey research methods* (pp. 520–521). SAGE Publications, Inc. http://dx.doi.org/10.4135/9781412963947.n333

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE Publications, Inc.

Mark, M. M. (2008). Building a better evidence base for evaluation theory: Beyond general calls to a framework of types of research on evaluation. In N. L. Smith & P. R. Brandon (Eds.), *Fundamental issues in evaluation* (pp. 111–134). Guilford Press.

Miles, M. B., Huberman, A. M., & Saldana, J. (2020). *Qualitative data analysis: A method sourcebook* (4th ed.). SAGE Publications, Inc.

Miller, R. L., & Campbell, R. (2006). Taking stock of empowerment evaluation: An empirical review. *American Journal of Evaluation, 27*(3), 296–319. https://doi.org/10.1177/1098214006027003 03

Moreau, K. A., & Cousins, J. B. (2012). Program evaluation in family-centred pediatric rehabilitation settings: A review of evaluation studies and the potential use of participatory and collaborative evaluation approaches. *Evaluation Journal of Australasia,11*(2), 3–13. https://doi.org/10.1177/1035719X1101100202

Patton, M. Q. (1997). *Utilization-focused evaluation: The new century text*. SAGE Publications, Inc.

Patton, M. Q. (2005). Utilization-focused evaluation. In S. Mathison (Ed.), *Encyclopaedia of evaluation* (pp. 429–432). SAGE Publications, Inc. https://dx.doi.org/10.4135/9781412950558

Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed.). SAGE Publications, Inc.

Samuelson, B. L., Smith, R., Stevenson, E., & Ryan, C. (2013). A case study of youth participatory evaluation in co-curricular service learning. *Journal of the Scholarship of Teaching and Learning*, *13*(3), 63–81.

Seidman, I. (2013). *Interviewing as qualitative research: A guide for researchers in education and the social sciences* (4th ed.). Teachers College Press.

Sheldon, J. F. (2016). *Evaluation as social intervention: An empirical study of empowerment evaluation practice and principle effects on psychological empowerment and self-determination outcomes* [Unpublished doctoral dissertation]. The Claremont Graduate University.

Smith, N. L. (1993). Improving evaluation theory through the empirical study of evaluation practice. *Evaluation Practice*, *14*(3), 237–242. https://doi.org/10.1177/109821409301400302

Upshur, C. C., & Barreto-Cortez, E. (1995). What is participatory evaluation (PE)? What are its roots. *The Evaluation Exchange, 1*(3/4), 1–7.

Varpio, L., Ajjawi, R., Monrouxe, L. V., O'Brien, B. C., & Rees, C. E. (2016). Shedding the cobra effect: Problematising thematic emergence, triangulation, saturation and member checking. *Medical education, 51*(1), 40–50. https://doi.org/10.1111/medu.13124

Wandersman, A., Alia, K. A., Cook, B., & Ramaswamy, R. (2015). Integrating empowerment evaluation and quality improvement to achieve healthcare improvement outcomes. *British Medical Journal*, *24*(10), 645–652. https://doi.org/10.1136/bmjqs-2014-003525

Wandersman, A., Imm, P., Chinman, M., & Kaftarian, S. (2000). Getting to outcomes: A results-based approach to accountability. *Evaluation and Program Planning, 23*(3), 389–395.

Wanzer, D. & Wisner, D. (2020, April 1). Research on evaluation: It takes a village (the problem). *Dana Wanzer*. https://danawanzer.com/roe-theproblem/

Zeller-Berkman, S., Muñoz-Proto, C., & Torre, M. E. (2015). A youth development approach to evaluation: Critical participatory action research. *Afterschool Matters*, *22*, 24–31.

https://files.eric.ed.gov/fulltext/EJ1083955.pdf

Appendix

Table A1. Table of Specifications for Phase 1 Survey

Dimension	Corresponding survey item numbers
Level of evaluator expertise and confidence in evaluating programs involving youth	Questions 1-3
Level of evaluator expertise and confidence in conducting EE of programs targeting youth	Question 4-5
Nature of EE conducted to evaluate programs targeting youth	Questions 6-17
Nature of evaluator experience in conducting evaluations of programs targeting youth	Questions 18-19

Table A2. Specifications for Phase 2 Interview Guides

Dimension	Corresponding interview questions
Context	Questions 1–3
Stakeholder characteristics	Questions 4–5
Organization characteristics	Question 6–7
Evaluator characteristics	Questions 8–9
Evaluation resources	Questions 10–11
External factors1	Question 12–13
Other factors	Question 14–16

Table A3. Phase 1 Number of Survey Respondents

Survey question subject area	Number of respondents for a given question	Number of respondents that elected to respond to question	Number of missing respondents
Survey respondents' type of discipline	108	67	41
Whether the respondent has evaluated programs involving youth	108	108	0
Number of programs involving youth evaluated by survey respondents, throughout their career	84	76	8
Number of programs involving youth evaluated by survey respondents, over the past year	84	75	9

Whether the respondent has used EE to evaluate programs involving youth	84	76	8
Reason why survey respondents used EE to evaluate programs involving youth	35	33	2
Evaluators' perceptions about the extent to which stakeholders were involved in particular EE activities	35	24	11
Survey respondents' confidence to use EE to evaluate programs targeting youth	84	75	9
Survey respondents' level of training in EE	84	74	10
Survey respondents'' rating of their understanding of EE	84	75	9
Survey respondents' experience conducting research on EE	84	74	10
Reason why survey respondents did not use EE to evaluate programs involving youth	41	41	0
Type of evaluation approach used by survey respondents who do not use EE to evaluate programs involving youth	41	41	0