

An Exploratory Study of the Factors Influencing the Relationship between
Program Evaluators and Their Youth Grantee Clients

by Mary Biehusen Podmostko

A.B., 1973, Smith College
M.A., 1997, The George Washington University

A Dissertation Submitted to

The Faculty of
The Graduate School of Education and Human Development
of The George Washington University
in partial fulfillment of the requirements
for the degree of Doctor of Education

August 31, 2009

Dissertation directed by:

Pamela J. Leconte
Assistant Research Professor of Special Education

The Graduate School of Education and Human Development of The George Washington University certifies that Mary Biehusen Podmostko has passed the Final Examination for the degree of Doctor of Education as of July 28, 2009. This is the final and approved form of the dissertation.

An Exploratory Study of the Factors Influencing the Relationship between
Program Evaluators and Their Youth Grantee Clients

Mary Biehusen Podmostko

Dissertation Research Committee

Pamela J. Leconte, Assistant Research Professor of Special Education,

Dissertation Director

Carol Kochhar-Bryant, Professor of Special Education, Committee Member

Sonia Jurich, Senior Research Associate, RMC Research Corporation, Committee

Member

Dedication

To the two people who never gave up on me:

my husband, Charles Podmostko,

and my advisor and dissertation director, Pam Leconte.

I wouldn't have made it this far without either one of you.

Acknowledgments

I would like to thank my academic advisor and dissertation chair, Dr. Pam Leconte, for her commitment and support during my doctoral program. Dr. Leconte's belief in me never wavered through my personal and professional transitions as well as my challenges in developing a viable dissertation study. I will always appreciate her support and guidance and look forward to working with her in the future.

To my committee members, Dr. Carol Kochhar-Bryant and Dr. Sonia Jurich: Thank you for your insights, challenges, and support during my coursework and dissertation.

To my external reviewers, Dr. Salvatore Paratore and Dr. George Tilson: Thank you for sharing the knowledge and insights that enriched and improved my dissertation.

To my friend, Marlene Chait, for your support and inspiring example of never giving up in the most challenging circumstances.

To Nancy Gilmore, Director of Doctoral Student Services, for regularly going above and beyond the call of duty to keep me on track.

To the Post Masters Appeals Committee, for believing in me and my research and for giving me the chance to prove that GWU's investment was worth it.

Abstract of Dissertation

An Exploratory Study of the Factors Influencing the Relationship between Program Evaluators and their Youth Grantee Clients

This study was designed to explore factors influencing the evaluator-client relationship in recent evaluations of programs serving transition age youth (ages 14-25) from the perspective of professional evaluators working with youth grantees.

Study data were collected by a researcher-developed electronic survey of 396 members of the American Evaluation Association who self-identified as providing evaluation services to youth programs serving the targeted age group. Both quantitative and qualitative were collected. Descriptive data were collected on the youth organization, the evaluator or evaluation consultant, and factors that were identified in the evaluation literature as characteristics, challenges, or threats influencing evaluation implementation or validity.

The study found that survey respondents worked with a variety of organizations providing youth services and had multiple roles as evaluators. Most respondents reported some level of staff or stakeholder anxiety related to the evaluation. Outcome/impact evaluation was the most common type of evaluation. A majority of respondents reported evaluation challenges related to inadequate or problematic data; lack of staff resources such as time or funding; lack of staff cooperation, collaboration, or communication; rigid organizational procedures or settings; limited methodological capacity of staff; or threats to validity related to mutating interventions and outside forces. The majority of respondents also reported evidence of successful evaluation implementation and usefulness. Qualitative responses linked identified challenges to the evaluator-client relationship. Resources used by evaluators spanned a wide range of content and delivery

methods. Examination of the data generated a number of questions regarding the evaluator-client relationship which, when investigated, will be instrumental in filling gaps in the literature on this important relationship.

Table of Contents

Dedication.....	iii
Acknowledgments.....	iv
Abstract of Dissertation.....	v
Table of Contents.....	vii
List of Figures.....	x
List of Tables.....	xi
CHAPTER ONE	
Overview of the Study.....	1
Statement of the Problem.....	2
Statement of Potential Significance.....	7
Theoretical Foundation and Conceptual Framework.....	8
Research Purpose and Research Questions.....	10
Summary of the Methodology.....	11
Delimitations, Limitations, and Assumptions.....	13
<i>Delimitations</i>	13
<i>Limitations</i>	14
<i>Assumptions</i>	15
Definitions and Abbreviations.....	16
CHAPTER TWO	
Introduction: Topics, Purposes, and Methods of the Literature Review.....	19
Theoretical Framework.....	20
<i>Program Evaluation Theory</i>	21
<i>Organizational Change Theory</i>	24
<i>Constructivist Theory</i>	27
Current Context for Youth Programs and Initiatives.....	29
<i>Government Calls for Improved Program Performance and Evaluation</i>	30
<i>Private Sector Calls for Improved Program Performance and Evaluation</i>	38
<i>Funding Programs and Services for Transition-Age Youth</i>	41
<i>Summary</i>	46
Performance of Programs and Initiatives Serving Transition-Age Youth.....	47
<i>Performance of Higher Education</i>	48
<i>Performance of K-12 Education Systems</i>	50
<i>Performance of “Second Chance” Programs</i>	55
<i>Performance of Discretionary Grant Funded Organizations</i>	57
<i>Summary</i>	60
The Challenges of Youth Program Evaluation.....	60
<i>Summary</i>	67

The Practice of Evaluating Youth Programs and Services.....	67
<i>Evaluation Resources</i>	69
<i>Youth Participatory Evaluation</i>	74
<i>The Evaluation Plan</i>	76
<i>The Evaluator-Client Relationship</i>	85
<i>Factors Influencing the Evaluator-Client Relationship</i>	90
<i>Summary</i>	94
Conceptual Framework for the Research Study.....	95
Inferences for the Forthcoming Study.....	97

CHAPTER THREE

Research Purpose and Questions	100
Overview of Methodology.....	101
Research Design.....	102
Population and Sample.....	104
Instrumentation.....	104
Data Collection Procedures.....	107
<i>Sampling Design and Procedures</i>	109
<i>Instrument and Tools Used or Designed</i>	111
<i>Validity and Reliability</i>	111
<i>Pre-Testing and Piloting Issues</i>	112
Data Processing and Procedures of Data Analysis.....	112
Human Participants and Ethics Precautions.....	114
Research Design Limitations.....	115

CHAPTER FOUR

Research Question One.....	117
<i>Type of Youth Organization Evaluated</i>	117
<i>Primary Funding Source for Youth Grant Program Evaluated</i>	118
<i>Primary Youth Target of Grant Program</i>	119
<i>Primary Youth Outcome of Grant Program</i>	121
<i>Primary Reason Youth Organization Hired an Evaluator</i>	122
Research Question Two.....	123
<i>Type of Organization for Which Evaluator Worked</i>	123
<i>Relationship of Evaluator to Youth Organization</i>	124
<i>Staff or Stakeholder Anxiety Encountered by Evaluator</i>	126
Research Question Three.....	128
<i>Primary Focus of Evaluation</i>	128
<i>Evaluation Planning Process Components</i>	130
<i>Youth Involvement in Evaluation Planning and Implementation</i>	131
<i>Challenges to Implementation of the Evaluation Plan</i>	133
<i>Threats to Validity of the Evaluation</i>	134
<i>Evidence of Successful Implementation and Usefulness</i>	136
<i>Audiences with Whom Evaluation Results Were Shared</i>	138

Qualitative Data Analysis.....	140
<i>Implementation Challenges or Threats</i>	140
<i>Resources for Youth Organization Evaluations</i>	144
<i>Open Comment Responses</i>	147
Instrument Reliability.....	151
Summary of Results.....	151
Conclusion.....	154
 CHAPTER FIVE	
Summary of Findings.....	155
<i>Research Question One: Characteristics of Youth Organizations</i>	155
<i>Research Question Two: Characteristics of Evaluators</i>	156
<i>Research Question Three: Evaluation Planning and</i> <i>Implementation Characteristics</i>	156
Questions Generated by Findings.....	158
Recommendations for Further Research.....	162
<i>Refinement and Validation of the Research Instrument</i>	163
<i>Complementary Surveys</i>	164
<i>In-Depth Investigations</i>	165
Implications for Programs Serving Youth with Disabilities.....	166
Recommendations for Policy and Practice	168
Recommendations for Professional Development	169
References.....	172
Appendix A: Survey Cover E-mail and Survey Instrument.....	203
Appendix B: Supporting Literature for Survey Questions.....	213
Appendix C: Pilot Protocol.....	218
Appendix D: Summary Results of the Pilot Survey.....	220
Appendix E: Summary of Question 17 Responses	230

List of Figures

1. Innovation Network Logic Model for Possible ODEP Grantee.....	81
2. Parallel Universe Conceptual Framework for Research Study	96
3. Flow Chart for Methodology	103

List of Tables

1. GAO Reports on Program Performance and Evaluation.....	32
2. A Sampling of Program Evaluation Resources.....	71
3. Possible Evaluation Planning Questions using the RUFDATA Model and ODEP SGA Guidelines.....	78
4. Calendar of Events for Implementing the Study.....	108
5. Type of Youth Organization Evaluated.....	118
6. Funding Source for Youth Grant Program.....	119
7. Youth Target.....	120
8. Youth Outcome.....	121
9. Reason for Hiring an Evaluator.....	122
10. Type of Evaluator Organization.....	123
11. Evaluator Relationship to Youth Organization.....	125
12. Staff/Stakeholder Anxiety Encountered.....	127
13. Focus of Evaluation.....	129
14. Evaluation Planning Process Components.....	130
15. Youth Involvement in Evaluation.....	132
16. Challenges to Implementation.....	133
17. Threats to Validity.....	135
18. Evidence of Success.....	137
19. Audiences for Evaluation Results.....	139
20. Resolution of Implementation Challenges or Threats.....	142

21. Evaluator Resources Summary.....	145
22. Evaluator Comments.....	148
23. Internal Consistency/Reliability of Instrument	151

CHAPTER ONE

Overview of the Study

This study explored factors influencing the relationship between program evaluators and their youth grantee clients who were funded by a broad range of competitive, discretionary grants with the goal of improving services and outcomes for transitioning youth ages 14-25. The study is grounded in (a) the American Evaluation Association's *Guiding Principles for Evaluators* (2004); (b) The Joint Committee on Standards for Educational Evaluation's *Program Evaluation Standards* (1994); and (c) the work of Carman (2007); Thayer (2006); Cecil (2004); Jackson (2004 and n.d.); Weiss (1998a & b); Patton (1997); Tilson (1986); and others.

Historically, programs and systems serving transition-age youth have had problems documenting their effectiveness (ACT, Inc., 2007; Economic and Development Review Committee, 2007; Employment and Training Administration, 2007; New Commission on the Skills of the American Workforce, 2007; What Works Clearinghouse, 2007; Horn, 2006; National Center for Public Policy and Higher Education, 2006; Barton, 2005; Carey, 2004; Gandal, Slover, Kraman, & Pache, 2004; Partnership for 21st Century Skills, 2004; Gladieux & Swail, 1998). Evaluations that “seek to answer the question: Are participants gaining the benefits that they were intended to receive? Or...what is happening to them because of the program's intervention?” (Weiss, 1998a, p. 5) are often required by program funders or sponsors (Rossi, Lipsey & Freeman, 2004)—and the consequences of a negative evaluation can be severe. For example, performance indicators and evaluation results are currently being used to dictate the use of federal stimulus funds (Samuels, 2009; U.S. Department of

Education, 2009, April 13) and end funding of programs with poor evaluation results (Klein, 2009; U.S. Department of Education, 2009, May 7).

This study explored the factors influencing the relationship between evaluators and youth grantees serving youth ages 14-25 by (a) describing the characteristics of evaluators, grantees, and the evaluation process that impact the evaluator/client relationship, (b) identifying areas for further study, (c) discussing implications for evaluation policy and practice, and (d) generating questions about the resulting data. A survey was conducted of members of the American Evaluation Association who served as evaluators or evaluation consultants for youth organizations funded by private or government entities to improve outcomes for youth ages 14-25. Data collected included descriptive information on the type of youth organization and program and a number of factors that impact the evaluator-client relationship such as the role of the evaluator, the evaluation focus, youth participation in the evaluation process, challenges to evaluation implementation, threats to evaluation validity, and evidence of successful evaluation implementation. A broad definition of youth program/youth service organization is employed in this study and includes youth services related to education, careers, employment, healthy choices, youth leadership, and independent living.

Statement of the Problem

In discussing evaluation design, Patton (1990, p. 162) said “...*there are no perfect research designs*. There are always trade-offs. These trade-offs are necessitated by limited resources, limited time, and limits on the human ability to grasp the complex nature of social reality.” Problems with the design and implementation of program evaluation plans have been documented in a number of contexts (Howell & Yemane,

2006; Jaycox, McCaffrey, Ocampo, Shelley, Blake, Peterson, Richmond & Kub, 2006; Renger, 2006). The expectations for evaluations vary among program providers such as community-based organizations, school-based entities, institutions of higher education, and others (Weiss, 1998b; Braskamp, Brandenburg & Ory, 1987; Tilson, 1986).

Many program providers have had limited knowledge of or formal training in designing and implementing program evaluations (Stoecker, 2007; Peterson & Randall, 2006; Mulroy & Lauber, 2004; Donaldson, Gooler, & Scriven, 2002; Tilson, 1986). Program providers may also have different attitudes and understandings of the purposes of evaluation (Carman, 2007; Tilson, 1986). Evaluation difficulties may surface in the original grant proposal, during program implementation, or in the final reporting process. As a result, the effectiveness of the evaluation may be impaired and the desired results may not be attained and/or documented.

A number of studies and reviews have found mixed results for programs serving transition age youth (Economic and Development Review Committee, 2007; New Commission on the Skills of the American Workforce, 2007; Branch-Smith, Gray, Fruchter, Hernandez, Joselowsky, Nichols-Solomon, Simmons, & Warren, 2006; Horn, 2006; National Center for Public Policy and Higher Education, 2006; Wagner, Newman, Cameto, Levine, & Garza, 2006; Barton, 2005; Carey, 2004; Bain, 2003; Blank, Melaville, & Shah, 2003; Gladieux & Swail, 1998). Determining program results can be difficult for a number of reasons including data or study design challenges (Chudowsky, Chudowsky & Kober, 2007; What Works Clearing house, 2007; Curran, 2005; National Institute of Statistic Sciences & Education Statistics Services Institute, 2005; Sum,

Harrington, Bartishevich, Fogg, Khatiwada, Motroni, Palma, Pond, Tobar, & Trub'sky, 2003).

A 2007 What Works Clearinghouse (WWC) review of school- and community-based dropout prevention programs is one example of the challenges in determining program effectiveness of federally-funded youth programs. WWC reviewed 58 studies of 16 programs and found 11 programs with studies that met WWC evidence standards with or without reservations. Three domains of interventions were identified: Staying in School, Completing School, and Progressing in School. Of the 11 programs meeting WWC standards, only one was found to have an unqualified rating of “positive effects” for a tutoring intervention in the “Staying in School” domain (WWC, 2007, p. 7).

Difficulties encountered by the WWC included defining the effects operationally and then determining if they were positive, that is, that they provided “strong evidence with no overriding contrary evidence” (WWC, 2007, p. 2). For example, one study showed a greater range of improvement in the “Staying in School” domain than another but had a lower rating due to differences in the “quality of the research design...statistical significance of findings...and consistency in findings across studies” (WWC, 2007, p. 5).

The client-evaluator relationship is a construct with both procedural and interpersonal components, both of which are critical for planning and implementing effective evaluations, but the literature is limited on specific aspects of the practice of evaluation. Henry and Mark (2003, quoted in Carman, 2007, p. 61) noted that “We know remarkably little about how evaluation is being practiced, why it is being practiced, by whom, and to what effect....Because there is a serious shortage of rigorous systematic evidence.” The literature does show that members of the American Evaluation

Association who provide program evaluation services are guided by two documents that describe guidelines for professional practice, *American Evaluation Association Guiding Principles* (AEA, 2004) and *The Program Evaluation Standards* of The Joint Committee on Standards for Educational Evaluation (1994). These documents focus primarily on structural or procedural issues such as negotiations, disclosure, cultural and disability competence, conflict prevention and resolution, professional ethics, and standards of professional practice.

The literature also identifies several factors that influence the evaluator-client relationship, although few are discussed in the context of that relationship. For example, evaluator roles are described as

- judge of program merit or worth (Scriven, 1991; Stufflebeam, 1994; Patton, 1997);
- auditor, inspector, or compliance/accountability evaluator (Barkdoll, 1980; Patton, 1997);
- researcher/scientist (Barkdoll, 1980; Weiss, 1998a; Patton, 1997);
- evaluation facilitator or guide (Weiss, 1998a; Patton, 1997; Wiltz, 2005);
- team member with evaluation perspective or internal advisor; (Patton, 1997, Wiltz, 2005);
- collaborator or partner (Barkdoll, 1980; Weiss, 1998a; Patton, 1997; Wiltz, 2005);
- critical friend (Rallis & Rossman, 2000; Weiss, 1998a);
- program improvement consultant (Patton, 1997);
- educator or capacity builder (Morabito, 2002; Wiltz, 2005; Weiss, 1998a);

- organizational development consultant or change agent (Owen, Lambert & Stringer, 1994; McClintock, 2003; Wiltz, 2005);
- empowerment facilitator who trains program administrators or staff to conduct evaluations on their own (Weiss, 1998a; Patton, 1997);
- supporter of a cause or advocate for social change (Patton, 1997; Morabito, 2002);
or
- program advocate (Wiltz, 2005).

Other factors that must be considered or managed in the evaluation planning and implementation process will influence the evaluator-client relationship but, again, are not described in the context of that relationship. The factors include the reasons for hiring an evaluator or evaluation consultant (Kegeles, Rebchook, & Tebbetts, 2005; Rossi, Lipsey & Freeman, 2004; Manela & Moxley, 1999; Thayer, 2006), the focus of the evaluation (Hanson, 1998) the evaluation plan and its components (Howell & Yemane, 2006; Jackson, 2004), challenges to evaluation implementation (Taut & Alkin, 2003; Weiss, 1998a; Cecil, 2004; Donaldson, Gooler & Scriven, 2002; Kegeles, Rebchook, & Tebbetts, 2005; Stoecker, 2007; Hasenfeld & English, 1974; Gilliam, Davis, Barrington, Lacson, Uhl, & Phoenix, 2002); threats to the evaluation validity (Jackson, n.d; Taut & Alkin, 2003; Weiss, 1998a; Mulroy & Lauber, 2004, Peterson & Randall, 2006), reports of the usefulness of the evaluation (Cecil, 2004: Gilliam, Davis, Barrington, Lacson, Uhl & Phoenix, 2002, Thayer, 2006); and audiences with whom evaluation results may be shared (Weiss, 1998a). This study is a first step in filling the gap in the literature on the subject of factors influencing the evaluator-client relationship.

Statement of Potential Significance

The stakes are high for providers of youth services as funders, both governmental and private, are taking a hard look at how effectively their funding is used (Klein, 2009; Samuels, 2009; U.S. Department of Education, 2009, April 13; U.S. Department of Education, 2009, May 7; The U.S Office of Management and Budget, n.d.a; Barton & Wilhelm, 2007; Buechel, Keating & Miller, 2007; Orosz, 2001) and how to more effectively address “difficult social problems” (Buechel, Keating, & Miller, 2007, p. 5). The stakes are especially high for at-risk groups such as students with disabilities who have lower rates of employment and post-secondary education enrollment than students without disabilities (Wagner, M., Newman, L., Cameto, R., Levine, P., & Garza, N., 2006).

This study explored factors influencing the evaluator-client relationship for evaluations conducted with youth program operators in a broad range of service environments (e.g., nonprofits; community-based organizations; public schools, special education transition programs, ABE/GED programs, colleges and universities, and One-Stop youth programs). It is hoped that questions generated by the study will lead to further investigation that will ultimately build the capacity of youth program operators to become better consumers of evaluation services, or at least better assess the quality of their evaluations, inform the practice of evaluators working with youth organizations and similar entities, and assist funders and policymakers in developing guidelines and incentives supporting effective evaluation for youth programs and grants.

Theoretical Foundation and Conceptual Framework

The theoretical foundation of the study is drawn from program evaluation theory, organizational change theory, and constructivist theory. Program evaluation theory combines evaluation theory and program theory (Patton, 2008) and explains what happens (or is supposed to happen) as a result of the program intervention. Logic models, theories of change, theories of action, impact pathways analysis, and outcomes mapping are examples of types of program theory (Patton, 2008). Program evaluation may be defined as “the *systematic assessment* of the *operation* and/or the *outcomes* of a program or policy, compared to a set of *explicit* or *implicit standards*, as a means of contributing to the *improvement* of the program or policy” (Weiss, 1998a, p. 4). In other words, “a program is a theory and evaluation is its test” (Rein, 1981, quoted in Weiss, 1998a, p. 55).

Organizational change theory is based on systems thinking which describes how people behave in organizations (Carson Research Consulting, 2007). Basic principles of systems thinking are (a) change is slow, (b) cause and effect may take a long time to surface, and (c) systems are complex (Carson Research Consulting, 2007). People resist change, and many organizational change efforts fail as a result, regardless of whether the organization is a business (Kotter, 2006), a school system (Hord, Rutherford, Huling-Austin, & Hall, 1989) or another type of organization (Carson Research Consulting, 2007; Wons, 1999).

Planning and implementing a program evaluation often results in significant organizational change for an organization, which may result in fear and defensive or negative work behaviors (Wheatley, 2002). Making the evaluation or change process

work requires understanding that organizations are “dense network[s] of relationships” (Wheatley, 2006) and comprehending the importance of those relationships (Wheatley, 2006). The presence of an evaluator in the organization creates a new and potentially stressful web of relationships with staff and stakeholders. Establishing trust (The Barefoot Collective, 2009) and understanding that “people must be engaged in meaningful work together if they are to transcend individual concerns and develop new capacities” (Wheatley, 2002, p. 3) will be part of the evaluator’s responsibilities.

Constructivist theory “explains how knowledge is produced in the world, as well as how students learn” (Gordon, 2009). Neal (2004, p. 6) cites Draper (2002) and Laroche and Bednarz (1998) in describing constructivism as “the philosophy or belief that learners create their own knowledge based on interactions with their environment including their interaction with other people...experience...environment... [and] language.” Constructivism also explains that people resist changing their minds when they encounter new information (Hyslop-Margison & Strobel, 2008, pp.78-79) and that “the outcomes of the learning process are varied and often unpredictable” (Walker & Lambert, 1995, p. 18). The result of constructivist thought processes is that there are “multiple perspectives and multiple truths” (Boghossian, 2006, p. 722) among different people in an organization.

The conceptual framework for the study is based on the theoretical frameworks of program evaluation, organizational change, and constructivism and descriptions in the literature of the disparate ways program managers and professional evaluators view and experience evaluation (Thayer, 2006; Taut & Alkin, 2003; Weiss, 1998b; Braskamp, Brandenburg, & Ory, 1987; Tilson, 1986). Differences in perspectives and understanding

may form “parallel universes” (a researcher-constructed concept) where one universe is inhabited by program staff and stakeholders and the other by evaluators. External forces and internal organizational and interpersonal dynamics impact the evaluator/client relationship, which is the bridge between the two universes. Internal and external forces may influence the evaluator-client relationship—positively by moving the parallel universes closer together or negatively by pushing them apart—thereby impacting the effectiveness of the evaluation.

Research Purpose and Research Questions

The literature suggests that a well-designed and implemented evaluation (Howell & Yemane, 2006) and effective evaluator-client relationships (Malloy & Yee, 2006) are necessary for effective evaluation. Therefore, the purpose of the study was to explore factors affecting the evaluator-client relationship in recent evaluations of programs serving transition age youth (ages 14-25) from the perspective of professional evaluators working with the youth grantees and to generate questions about the findings. The research questions for this study were:

1. What are the characteristics of the organizations whose youth program was evaluated?
2. What are the characteristics of the role of the evaluators or evaluation consultants?
3. What are the characteristics of program evaluation planning and implementation that may impact the evaluator/client relationship?
4. What questions relating to the evaluator-client relationship result from analyzing the data collected in questions 1, 2, and 3?

Summary of the Methodology

“Exploratory data analysis can be defined as the examination of data with minimal preconceptions about its structure through which it is hoped that relationships and patterns, at least some of which are unanticipated, will be discovered” (Wainer, 1978, p. 151). In the process, questions or hypotheses will be generated for future analysis (Tukey, 1980; Brilliner, 2002). Consequently, descriptive data was collected from professional evaluators who were members of the American Evaluation Association (AEA) and who had been involved with a broad range of grants serving youth in the 14 to 25 year age range. A researcher-constructed, electronic survey containing quantitative and qualitative data elements was used. Survey responses were analyzed to identify issues and patterns responding to the research questions. Professional evaluators were surveyed for their observations on the practice of youth program evaluation because they have been trained in evaluation, are more comfortable with evaluation terminology, data, and reporting, and, in the case of AEA members, have a track record of responding to surveys on evaluation practice.

A survey questionnaire was developed based on the research questions described above. The survey was sent to five experienced researchers and evaluators for review and was revised extensively based on the resulting input. After Institutional Review Board and American Evaluation Association review, the survey was piloted with 19 experienced evaluators recruited from the Southeast Evaluation Association, an affiliate of AEA, and EVALTALK, AEA’s professional listserv. This two-step process was used to determine the internal consistency of quantitative survey questions and validity of

constructs as well as identify inconsistencies, ensure ease of use of the electronic survey instrument, and correct problems with wording, instructions, clarity, and logic.

AEA members working in the United States who offer evaluation services to the community were invited to respond to the survey in accordance with Institutional Review Board (IRB) and AEA guidelines. The American Evaluation Association's 38 Topical Interest Groups (TIGs) were reviewed in order to identify likely sources of survey candidates. AEA members were selected for the survey because of their high professional standards, broad range of evaluation experiences, variety of evaluation positions (ranging from independent consultants to governmental evaluators), demonstrated concerns for promoting valid, responsive evaluation practice, and access to the Internet. It is also anticipated that certain types of sampling errors, such as central tendency and social desirability biases, will be greatly reduced because of AEA members' familiarity with evaluation concepts and methodologies.

A survey methodology was selected because the study was exploratory in nature. Data about the relationship between the evaluator and the grantee organization were collected to provide preliminary information about challenges and successes in the evaluation of youth programs and to identify issues for further investigation. An on-line survey was selected because it meets the criteria described in Ritter and Sue's (2007) summary of the advantages and disadvantages of on-line surveys:

Online surveys are relatively inexpensive to conduct, have the potential to collect a lot of data quickly, and can reduce overall survey error, because the data entry chore is eliminated. On the other hand, Internet-based surveys are practical and sensible only if the target respondents have access to computer technology and the researcher has access to a list of the potential respondents. (p. 5)

Emerging research on web-based surveys appears to indicate that responses to properly designed surveys with adequate invitation and follow-up procedures generate faster responses and comparable to significantly higher response rates than surface mail questionnaires for samples similar to AEA members (technologically literate, with ready access to the internet, and geographically dispersed) (Mitra, Jain-Shukla, Robbins, Champion, & Durant, 2008; Ritter & Sue, 2007; Kiernan, Kiernan, Oyler & Gilles, 2005; Kaplowitz, Hadlock, & Levine, 2004). AEA member access to internet technology is manifested by popular association listservs such as EVALTALK and EVALJOBS and by TIG listservs such as those of the Collaborative, Participatory, and Empowerment Evaluation TIG; the Minority Issues in Evaluation TIG; and the Non-Governmental Organizations TIG.

Delimitations, Limitations, and Assumptions

Delimitations

“Delimitations are restrictions placed on a research study by the researcher to limit the extent of the study” (Peters, 2009, p. 12). The research study was limited to evaluators or evaluation consultants who self-identified as meeting the following participant profile and who voluntarily responded to the survey:

1. American Evaluation Association member
2. Provider of evaluation services in the United States
3. Experience as an evaluator or evaluation consultant (a) within the last five years and (b) to a grant-funded organization serving youth ages 14-25

The experimental design, an exploratory survey of a purposeful sample of American Evaluation members, further limited generalizing to a larger population of evaluators.

Limitations

“Limitations are potential weaknesses or problems with the study that are identified by the researcher” (Creswell, 2002, quoted in Peters, 2009, p. 12). This study collected data from members of the American Evaluation Association who worked with youth organizations funded by grants to improve youth outcomes for transition-age youth (ages 14 to 25). However it was not known how many of the approximately 5,500 AEA members met the survey criteria. The resulting study was descriptive, non-experimental, and exploratory. A second limitation was that responses were based on participant perceptions of the questions, or, as one of the respondents stated, the questions were “subjective.” A third limitation was that the study invited participants from one professional evaluation association due to limited resources and convenience.

A fourth limitation was the number of factors that may have caused a low survey response rate, aside from the issue of whether the recipient met the survey criteria. Feedback from a few pilot and study participants described lost survey invitations, technical difficulties such as the recipient’s firewall settings, and external blocks on SurveyMonkey access. The timing of the survey, which was disseminated during the December holiday season, was also a factor, particularly for the 29% of AEA members were employed by colleges and universities (Goodman Research Group, 2008) that customarily close for several weeks during the holiday season. Additionally, survey fatigue of potential recipients may have affected response rates given that the study survey was scheduled between two other approved surveys of the AEA membership and followed a major membership scan conducted in 2007-2008.

Assumptions

Great care was taken by the researcher to eliminate errors in the handling of the data. Response errors were expected to be reduced because survey respondents were professional evaluators with some familiarity with survey methodology. Survey respondents were instructed to refer to their most recent evaluation in answering survey questions in order to reduce recall errors. Approximately 20-30% of the invited members were expected to meet the survey criteria based on information from a recent survey of American Evaluation Association members and because members of Topical Interest Groups with a high likelihood of serving youth programs were invited to take the survey. AEA survey data indicated that 95% of members were involved in program evaluation (Goodman Research Group, 2008, p.18) and that 37% worked in the content area of with K-12 education, 27% worked in youth development, and 26% worked in higher education (Goodman Research Group, 2008, p. 15).

A survey invitation was e-mailed to 3,892 AEA members. A second, follow up e-mail invitation was sent to 3,424 e-mail addresses from which no response (other than automatic out-of-office messages) had been received from the first survey invitation. Ultimately, 396 survey responses (a 10% response rate) were obtained from AEA members who self-identified as meeting the survey criteria. Survey results provided the perspective of one group—professional evaluators who were members of the American Evaluation Association—and may not represent the full range of AEA members or other evaluators working with youth grantees. However, their responses are valuable both for the insights they provide into the evaluator/client relationship and for the identification of areas for future research.

Definitions and Abbreviations

American Evaluation Association (AEA). “The American Evaluation Association is an international professional association of evaluators devoted to the application and exploration of program evaluation, personnel evaluation, technology, and many other forms of evaluation...” It has approximately 5500 members. (AEA web site, “About Us” page).

Constructivism. “The philosophical position that truth is contingent and conditional and that there are multiple perspectives and multiple realities...[and] that people in different geographic, cultural, or social locations construe knowledge, truth, and relevance in different ways, each of them legitimate and worthy” (Weiss, 1998a, p. 328).

Evaluator. Someone who conducts “technically challenging” evaluations “with reasonable competence” in several areas including “basic qualitative and quantitative methodologies...legal constraints... ethical analysis...conceptual geography...[and] evaluation-specific report design, construction, and presentation” (Scriven, 1996, pp.159-160).

Evaluation consultant. Someone who “helps practitioners understand the types of evaluative decisions they face and enhances[s] their decisions to deliberate well” (Schwandt, 2005, p. 99) or in a broader sense, someone with more technical knowledge and skill in evaluation than the client who hired him or her (Scriven, 1996, p. 159)

Evaluation sponsor or funder. “The person, group, or organization that requests or requires the evaluation and provides the resources to conduct it” (Rossi, Lipsey & Freeman, 2004, p. 426).

GPRA. Public Law 103-62, the Government Performance and Results Act of 1993. Law passed to improve performance, effectiveness and accountability through reforms focused on goal-setting, results, service quality, customer satisfaction, and public reporting and accountability.

OECD. The Organisation for Economic Co-Operation and Development, “a unique forum where the governments of 30 market democracies work together to address the economic, social, and governance challenges of globalisation as well as to exploit its opportunities” (OECD, 2006, p. 7).

Program evaluation. “The *systematic assessment* of the *operation* and/or the *outcomes* of a program or policy, compared to a set of *explicit* or *implicit standards*, as a means of contributing to the *improvement* of the program or policy” (Weiss, 1998a, p. 4).

Program theory. “The set of assumptions about the manner in which a program relates to the social benefits it is expected to produce and the strategy and tactics the program has adopted to achieve its goals and objectives” (Rossi, Lipsey & Freeman, 2004, p. 432).

Transition. Movement of youth “into the workplace, higher education, independent living, and other adult objectives” (Rothstein, 1990, p. 153).

Stakeholders. “Individuals, groups, or organizations having a significant interest in how well a program functions, for instance, those with decision-making authority over the program, funders and sponsors, administrators and personnel, and clients or intended beneficiaries” (Rossi, Lipsey & Freeman, 2004, p. 435).

Youth. Young people ages 14-25 who are transitioning to adulthood, including those with disabilities. (Timmons, Podmostko, Bremer, Lavin, & Wills, 2005, p. iii).

Youth programs/Youth service providers/Youth service organizations: Organizations, institutions, and agencies in a wide variety of settings at the community, state, and national levels that provide services to youth ages 14-25 in the areas of education and training; career preparation; work-based learning; youth development and leadership; activities connecting youth to transportation, health and mental health services, post-program supports, etc.; and/or family involvement and supports. (National Collaborative on Workforce and Disability for Youth, n.d.)

CHAPTER TWO

Introduction: Topics, Purposes, and Methods of the Literature Review

This literature review sets the stage for the research study by presenting a theoretical framework for the proposed study (program evaluation theory, organizational change theory, and constructivist theory) followed by an examination of the current reality of evaluating youth programs and initiatives based on (a) calls for change in accountability, performance, and evaluation of programs serving transition-age youth ages 14-25, (b) the performance of programs and initiatives serving transition age youth, and (c) challenges encountered in evaluating youth programs. The literature review then discusses aspects of the current practice of evaluating youth programs and services including the development and implementation of the evaluation plan and the evaluator-client relationship. The chapter ends with a conceptual framework and inferences for the proposed study.

The literature review examined the current literature regarding (a) theories of program evaluation, organizational change, and constructivism, (b) youth program performance and funding, and (b) program evaluation practice and challenges. Searches were conducted on several fronts. Targeted searches were conducted via ALADIN, ERIC, PsycINFO, Google, and Sage Publications. Current periodicals and news lists provided rich sources of applicable literature as well. Publications and resources related to youth programs and performance such as *Education Week's Research Connections*, *CEC SmartBrief*, and *GrantStation Insider* were reviewed regularly.

Publications and resources related to program evaluation included three commonly used evaluation references, two American Evaluation Association (AEA)

journals (*American Journal of Evaluation* and *New Directions for Evaluation*), the AEA e-newsletter, EVALTALK (the AEA listserv), and reports and technical assistance materials developed for program evaluations by funders and other entities. Government Accountability Office publications, dissertations, and research methodology texts and resources were also reviewed.

A literature review via ProQuest and other social sciences databases was conducted to determine if an existing survey instrument could be used for the research study and to identify survey methodology resources. No survey instrument was found that addressed the study's specific research questions.

Theoretical Framework

Program evaluations are conducted for a number of purposes. Many are conducted to determine which programs are effective (e.g., determine if programs are meeting their goals) or to revise them to make them more effective (Rossi, Lipsey & Freeman, 2004). They are also employed to make mid-course corrections; continue, cut, expand, or end a program; test a new program idea; select one of several program versions; or decide whether to continue a program (Weiss, 1998a). Program evaluation can also be conducted as part of organizational development or learning through recording program history, providing feedback to practitioners, highlighting program goals, demonstrating accountability, or studying social interventions. (McClintock, 2003; Weiss, 1998a). Unfortunately, some evaluations are conducted to postpone decisions, duck responsibility, or shore up public relations (Weiss, 1998a).

Evaluators and youth program staff and stakeholders vary greatly in their knowledge and understanding of program evaluation theory. They experience the

evaluation process from different places within or outside the organization being evaluated. They construct individual, unique meanings based on their different cultures, life experiences, education, and motivations. As a result, their understanding of program evaluation and the way they and their organizations react to and implement the changes required to design and implement an evaluation may be very different. Program evaluation theory, organizational change theory, and constructivist theory provide the theoretical underpinnings of this study and help explain the dynamics of the program evaluation process and the evaluator client/relationship.

Program Evaluation Theory

Program evaluation theory combines evaluation theory and program theory (Patton, 2008). Program theory explains what happens (or is supposed to happen) as a result of the program intervention. Logic models, theories of change, theories of action, impact pathways analysis, and outcomes mapping are examples of types of program theory (Patton, 2008, p. 336).

A logic model describes the sequence of events that is expected to occur in a program. “Specifying causal mechanisms transforms a logic model into a theory of change” (Patton, 2008, p. 336). The terms, theory of change and theory of action, are often used interchangeably and include the causal mechanism (intervention or action) that is expected to produce the desired effect. Patton concludes that

confusion reigns in the language describing how program activities lead to program outcomes... The best designation is the one that makes the most sense to primary intended uses [sic]—the term they resonate to and has meaning within their context given the intended uses of the evaluation” (2008, p. 339).

If program theory explains what happens during program interventions, it follows that evaluation theory describes what happens during evaluations. Program evaluations

are initiated for a number of reasons and purposes, and as a result there are many types of program evaluations with varying purposes and foci.

Program evaluation may be defined as “the *systematic assessment* of the *operation* and/or the *outcomes* of a program or policy, compared to a set of *explicit* or *implicit standards*, as a means of contributing to the *improvement* of the program or policy” (Weiss, 1998a, p. 4). Program evaluation may be used to “study, appraise, and help improve social programs, including the soundness of the programs’ diagnoses of the social problems they address, the way the programs are conceptualized and implemented, the outcomes they achieve, and their efficiency” (Rossi, Lipsey, & Freeman, 2004).

Rossi, Lipsey, & Freeman (2004) define four common types of program evaluations and their purposes, which may overlap or occur simultaneously, as follows:

1. Summative evaluation – “Evaluative activities undertaken to render a summary judgment on certain critical aspects of the program’s performance, for instance, to determine if specific goals and objectives were met.” (pp. 65)
2. Formative evaluation – “Evaluative activities undertaken to furnish information that will guide program improvement.” (p.63)
3. Process evaluation – “A form of program monitoring designed to determine whether the program is delivered as intended to the target recipients. Also known as implementation assessment.” (p. 64)
4. Outcome evaluation – “An evaluative study that answers questions about program outcomes and impact on the social conditions it is intended to ameliorate. Also known as an impact evaluation” (p. 63)

A “needs assessment that verifies and maps the extent of a problem” in order to determine client needs was identified by Hanson (1998, p. 44) as an additional type of evaluation.

The end results of the program as measured by an evaluation may be short-term or long-term. Outcomes as defined by the United Way (Hatry, van Houton, Plantz, & Taylor, 1996, p. 18), for example, are “benefits or changes to individuals or populations during or after participating in program activities. They are influenced by a program’s outputs. Outcomes may relate to behavior, skills, knowledge, attitudes, values, condition, status or other attributes.” Others may define outcomes as longer term impacts for the community (Weiss, 1998a, pp. 8 & 331) such as increased employment rates or reduction in crime rates.

Program evaluation theory describes how a program and its evaluation work. Weiss defines program theory as “the set of beliefs which underlie action” (1998a, p. 55) and program implementation theory as “what is required to translate objectives into on-going service delivery and program operation” (p. 58). She defines the program’s theory of change as the combination of program theory and implementation theory (p. 58).

Another way to look at program evaluation theory is that “a program is a theory and evaluation is its test” (Rein, 1981, quoted in Weiss, 1998a, p. 55). Rossi, Lipsey & Freeman (2004, p.44) describe program theory as the “plan of operation, the logic that connects its activities to the intended outcomes, and the rationale for why it does what it does.”

A program theory, which often includes a logic model is typically composed of inputs, activities or interventions, interim outcomes or outputs and desired end results or

outcomes (Weiss, 1998a, p. 62; Hatry, van Houten, Plantz, & Taylor, 1996) and can be represented as follows:

$$\text{Inputs} + \text{Activities/Interventions} = \text{Outputs} + \text{Outcomes}$$

There are two types of change involved in designing and implementing a program evaluation—the change that the program is trying to make in the lives of youth as described in the program’s theory of change or logic model and the organizational change required for the evaluation to take place.

Organizational Change Theory

Organizational contexts and the assumptions and needs of the people involved determine how people in an organization respond to change. Regardless of how the change process is described or the approach used to facilitate it, organizational change is not easy. It is a complex process that requires planning and hard work to implement successfully. If the change involved is a program evaluation, which practitioners may or may not welcome or understand, the change process needed to plan and implement the evaluation may be challenging.

Organizational change theory is based on systems thinking which describes how people behave in organizations (Carson Research Consulting, 2007). Basic principles of systems thinking are (a) change is slow, (b) cause and effect may take a long time to surface, and (c) systems are complex (Carson Research Consulting, 2007). People resist change, and many organizational change efforts fail as a result, regardless of whether the organization is a business (Kotter, 2006), a school system (Hord, Rutherford, Huling-Austin, & Hall, 1989), or other types of organizations (Carson Research Consulting, 2007; Wons, 1999).

A number of theorists and practitioners have developed models or descriptions of what happens (or should happen) during organizational change (Kotter, 2006; Wons, 1999; Hord, Rutherford, Huling-Austin, & Hall, 1989). The following models describe organizational change processes in different ways, but it is important to note that organizations are composed of people, and it is at the individual level that change actually occurs. Some models address the impact organizational change has on personnel more explicitly than others.

Kotter's "Best of Harvard Business Review" process contains eight phases that must be properly implemented—in the correct order—for organizational change to be effective. The steps are

1. Establishing a sense of urgency
2. Forming a powerful guiding coalition
3. Creating a vision
4. Communicating the vision
5. Empowering others to act on the vision
6. Planning for and creating short-term wins
7. Consolidating improvements and producing still more change
8. Institutionalizing new approaches (Kotter, 2006, p. 2)

Wons' (1999) Ethical Means-Based Model is a complex, continuous cycle of organizational development composed of four components representing on-going conflicts: Survival/Failure, Delegation/Control, Mission/Special Interest, and Reevaluation/Diminishing Returns. These conflicts exist regardless of whether a specific change is being introduced and determine whether the change will be successful.

Hord et al's Concerns-Based Adoption Model (CBAM) is client-centered and focuses on the people involved in the change initiative. Change facilitators are taught to identify the seven stages of concern (awareness, informational, personal, management, consequence, collaboration, refocusing (1989, p. 31) among the people involved in the change and to help them work through the stages as part of the change facilitation process. The CBAM model is based on six assumptions about change:

1. Change is a process, not an event.
2. Change is accomplished by individuals.
3. Change is a highly personal experience.
4. Change involves developmental growth.
5. Change is best understood in operational terms.
6. The focus of facilitation should be on individuals, innovations, and the context. (Hord et al, 1989, pp. 5-6).

Havelock & Hamilton (2004, p. xxix) identified a practical model with seven stages for guiding change in special education that has broad application to complex systems. The seven stages are

1. Care: What is the concern? Who cares? How much? How much should they care?
2. Relate: Whose concern is it? Who are the key stakeholders? How do we relate to them?
3. Examine: How do we define the concerns as a solvable problem?
4. Acquire: How do we get help? Who has resources? Solutions? How do we gain access to these resources?

5. Try: How do we pick the best solution? How do we put the elements together?
6. Extend: How do we translate into action? Get consensus? Widen the circle of users?
7. Renew: How do we internalize process to continue innovating?

Planning and implementing a program evaluation often results in significant organizational change for an organization and the people in it, which may result in fear and defensive or negative work behaviors (Wheatley, 2002). Making the evaluation or change process work requires understanding that organizations are “dense network[s] of relationships” (Wheatley, 2006, p.1) and comprehending the importance of those relationships (Wheatley, 2006).

The presence of an evaluator in the organization creates a new and potentially stressful web of relationships with staff and stakeholders. Establishing trust (The Barefoot Collective, 2009) and understanding that “people must be engaged in meaningful work together if they are to transcend individual concerns and develop new capacities” (Wheatley, 2002, p. 3) are part of the evaluator’s responsibilities.

Constructivist Theory

Even if program practitioners and evaluators had exactly the same training and experiences with program evaluation and organizational change, their understandings of and reactions to the evaluation process would very likely be different. Constructivist theory provides insight into why this difference occurs by “explain[ing] how knowledge is produced in the world, as well as how [people] learn” (Gordon, 2009). Neal (2004, p. 6) cites Draper (2002) and Larochelle and Bednarz (1998) in describing constructivism as “the philosophy or belief that learners create their own knowledge based on interactions

with their environment including their interaction with other people... experience... environment... [and] language.”

Gordon (2009, p. 40) describes “serious problems with the fragmented and incoherent character of the literature on constructivism” leading to “a lack of clarity” on the subject. Much of the available literature on constructivism focuses on teaching and learning, but the theory can easily be transposed to evaluation. Gordon also notes that genuine learning is active, not passive (p. 47), that knowledge is individually *and* socially constructed, and that “the authority of knowledge...rests heavily on the teachers’ [or evaluators’] own knowledge and experience” (p. 48).

Additional characteristics of constructivism are (a) prior knowledge and experience influence how people respond to new information, (b) people resist changing their minds when they encounter new information, (c) people “learn that their constructed beliefs do not necessarily qualify as knowledge, and [d] that knowledge emerges from sources other than their own individual cognition” (Hyslop-Margison & Strobel, 2008, pp.78-79).

Walker and Lambert (1995, pp. 17-19) describe seven principles of constructivist learning theory that include metacognition and reflection as essential parts of the learning process and the conclusion that “the outcomes of the learning process are varied and often unpredictable.” The result of constructivist thought processes is that there are “multiple perspectives and multiple truths” (Boghossian, 2006, p. 722) among different people in an organization.

Constructivism, therefore, explains why two people can observe the same event or read the same book and interpret what happened or what they read in very different ways.

It also explains that people with different backgrounds, training, motivation, world view, and any number of other characteristics will probably experience the process of program evaluation very differently—particularly if one is a professional evaluator and the other is a youth service practitioner. Therefore, since the authority of evaluation knowledge customarily resides with the evaluator or evaluation team, their responsibilities include coaching program staff and stakeholders through a learning process that fosters reflective practice, demonstrating sensitivity to previous knowledge constructions, and encouraging ownership, voice, and collaboration (Boghossian, 2006, p. 722).

Program evaluation theory, organizational change theory, and constructivist theory overlap to provide a framework that adds depth and understanding to the discussion of why evaluations work or don't work and why programs do or do not meet their goals. The area in which the theories and practice overlap is the area in which the evaluator and his or her organizational clients work together, thereby delineating the evaluator/client relationship as a critical element of success.

Current Context for Youth Programs and Initiatives

Youth programs today operate in a context that is characterized by increasing demands for accountability and performance and decreasing availability of funding. Both of these trends place increasing importance on program monitoring and evaluation. This section examines the performance of youth programs funded by the federal government and the private sector, resulting efforts to improve performance accountability, and the current environment for funding youth programs.

Government Calls for Improved Program Performance and Evaluation

Program performance is a matter of integrity in awarding grants, managing programs, and documenting the achievement of desired program goals. A number of entities are calling for more accountability, more evaluation, better performance data, and better results for their investments (U.S. Office of Management and Budget, n.d.a; Barton & Wilhelm, 2007; Buechel, Keating & Miller, 2007; Executive Office of the President of the United States, Office of Management and Budget, 2007; Davis, Uhl, Barrington, Rowel, Squiers, Sharp, & O'Brien, 2004) in order to tie continued funding to some type of performance on the part of grantees based on accountability measures and/or evaluation plans.

Under *Public Law 103-62, The Government Performance and Results Act of 1993* (GPRA), federal agencies and their grantees are held accountable for achieving measurable performance goals and objectives. The ExpectMore.gov website (U.S. Office of Management and Budget, n.d.a), which rates the effectiveness of federally-funded programs, is one example of the increased scrutiny experienced by federal programs and grantees. The ExpectMore.gov web site lists nearly 800 federally-funded programs, including those that are funded by competitive and formula grants. In February 2006, 28% of the programs were rated as “not performing”, that is, “ineffective” or “results not demonstrated.” In May 2007, 25% were rated as “not performing”. (Note: The “performing” rating included programs that were rated as “effective,” “moderately effective,” or “adequate.”)

In 2000, the General Accounting Office (now the Government Accountability Office) (GAO) noted that GPRA did not require agencies to conduct program evaluations

in order to determine if program goals were achieved and that “many agencies had not given sufficient attention to how program evaluations would be used in implementing GPRA and improving program performance” (Kingsbury, 2000, p. 4). Some agencies, however, were using evaluations to support GRPA goals in order to

1. determine why performance goals were or were not being met,
2. identify specific actions for improving program performance, and/or
3. identify problems with data collection such as data quality, availability, accuracy, completeness, and consistency (Kingsbury, 2000).

The GAO concluded:

Over the last several years, we have noted that, government-wide, agencies’ capability to gather and use performance information has posed a persistent challenge to making GPRA fully effective. . . . More recently . . . we noted our continued concern that many agencies lack the capability to undertake the program evaluations that are often needed to assess a federal program’s contributions to results where other influences may be at work. (Kingsbury, 2000, p. 18)

In the seven years since this statement, the GAO has identified a number of problems in evaluating or determining performance of youth-serving programs funded by federal agencies with concomitant recommendations for improvement. Some of these problems relate to the grant award process itself. Table 1 contains a sampling of those reports.

Table 1

GAO Reports on Youth Program Performance and Evaluation

Report	Problems	Recommendations
Ekstrand, L.E. (2001). <i>Juvenile Justice: OJJDP reporting requirements for discretionary and formula grantees and concerns about evaluation studies.</i> (GAO 02-23).	Variations in information reported and site implementation, lack of comparison groups, and data collection problems.	OJJDP should address comparison group and data collection problems and initiate interventions to address problems.
Ekstrand, L.E. & Rezmovik, E.L. (2003). <i>Justice outcome evaluations: Design and implementation of studies require more national Institute of Justice attention.</i> (GAO 03-1091).	Some studies were well designed and implemented but others had problems with implementation or serious methodological programs.	Develop appropriate strategies and corrective measures; ensure only effectively designed and implemented studies are funded.
Bellis, D.D. (2004). <i>Workforce Investment Act: Labor actions can help states improve quality of performance outcome data and delivery of youth services.</i> (GAO 04-308).	Challenges in <ul style="list-style-type: none"> - serving out-of-school youth - providing mentoring and follow-up services, and - establishing and using optional interim performance measures. 	<ul style="list-style-type: none"> - Coordinate efforts between Education and Labor - Increase guidance and technical assistance - Establish standard monitoring procedures for data validation efforts.
Shaul, M.S., Edwards, D., & Fucile, T. (2004). <i>Special education: Improved timeliness and better use of enforcement actions could strengthen education's monitoring system.</i> (GAO 04-879).	<ul style="list-style-type: none"> - Weak data elements used in monitoring - States noncompliant in providing student services and meeting procedural requirements. - Noncompliance cases take many years to be resolved. 	<ul style="list-style-type: none"> - Provide additional guidance - Expedite resolution of noncompliance - Impose firm, realistic deadlines and initiate compliance agreement proceedings for corrections.

Report	Problems	Recommendations
Robertson, R.E. (2005). <i>Better measures and monitoring could improve the performance of the VR program.</i> (GAO 05-865).	<ul style="list-style-type: none"> -Performance measures are not comprehensive and monitoring does not result in timely feedback. -Performance of certain key populations is not measured. -No censure of poor performance, reward for strong performance, or dissemination of best practices. 	<ul style="list-style-type: none"> - Reevaluate performance measures - Improve timeliness of performance data and feedback. - Apply consequences for failure. - Consider developing incentives for successful performance. - Develop means for disseminating best practices.
Shaul, M.S. (2005). <i>No Child Left Behind Act: Education could do more to help states better define graduation rates and improve knowledge about intervention strategies.</i> (GAO 05-879).	<ul style="list-style-type: none"> - States use different definitions of graduation rates - No consistent guidance on accounting for certain types of students - Concerns about data accuracy persist, primarily due to problems with student mobility (p. 4) - Monitoring efforts may be insufficient for states to provide accurate data - Few state interventions for increasing graduation rates have been evaluated - Little dissemination of effective intervention. 	<ul style="list-style-type: none"> - Make information available to all states on graduation calculation modifications - Assess the reliability of state graduation rate data before developing interim rate estimates - Establish a timetable for evaluating research on dropout interventions - Disseminate information on programs with demonstrated effectiveness.

Report	Problems	Recommendations
<p>Bellis, D.D. (2005). <i>Education's data management initiative: Significant progress made, but better planning needed to accomplish project goals.</i> (GAO 06-06).</p>	<p>Education's Performance-Based Data Management Initiative has problems including</p> <ul style="list-style-type: none"> - Inability to resolve data differences between program offices - Lack of baseline data - No process for resolving data differences - Inability of states to provide reliable data - No plans for addressing implementation obstacles, providing additional assistance to states, or meeting state expectations. 	<p>Department of Education should develop</p> <ul style="list-style-type: none"> - A strategy to help states improve their ability to provide quality data - A process for reconciling differences between program offices - A clear plan for completing Performance-Based Data Management Initiative
<p>Shaul, M.S. (2006). <i>Discretionary grants: Further tightening of Education's procedures for making awards could improve transparency and accountability.</i> (GAO 06-268).</p>	<ul style="list-style-type: none"> - Competition plans were not finalized prior to start of competitions - Grant files lacked documentation - Unsolicited proposals had no standard format - Extensive technical assistance was provided to some applicants but not others. 	<ul style="list-style-type: none"> - Develop a more systematic format for unsolicited proposals - Ensure all competition plans are finalized before competitions begin - Screen all applicants for compliance and ensure that audit issues are addressed before the award (p. 19) - Document all required reviews.
<p>Robertson, R.E. (2006). <i>Summary of a GAO conference: Helping California youths with disabilities transition to work or postsecondary education.</i> (GAO 06-759SP).</p>	<ul style="list-style-type: none"> - Youth with disabilities do not receive sufficient training in vocational preparation, life skills, and transition planning. - Limited coordination and differences in program structure prevented the seamless provision of services. 	<ul style="list-style-type: none"> - Begin transition planning at a younger age, create internship programs during high school, and bring in mentors from work - Designate days for teachers to coordinate with other programs - Adopt common materials and definitions if feasible.

Report	Problems	Recommendations
<p>Robertson, R.E., (2007). <i>Federal disability assistance: Stronger federal oversight could help multiple programs' accountability</i>. (GAO 07-236).</p>	<ul style="list-style-type: none"> -Individual programs could not fully assess performance due to mixing of funds, had mixed success in meeting goals, and/or could not report progress due to difficulties with definitions and measures. -Federal oversight was uneven with little guidance to ensure compliance or consistent interpretation of program requirements. 	<ul style="list-style-type: none"> - Provide more effective leadership by establishing performance goals, being more proactive in disseminating guidance, and strengthening monitoring - Assess goals and measures and continue to develop performance management system - Improve oversight procedures
<p>Ashby, C.M. (2007). <i>No Child Left Behind Act: Education actions may help improve implementation and evaluation of Supplemental Educational Services</i> (SES). (GAO 07-738T).</p>	<ul style="list-style-type: none"> - Notification of parents is not timely or effective - Provider curriculum not aligned with district instruction - Contracting and coordination of services is problematic - State monitoring of services is limited - Evaluating improvement of student achievement is difficult - Uneven implementation and compliance - Policy clarification and assistance is needed 	<ul style="list-style-type: none"> - Clarify guidance and provide additional assistance - Consider expanding pilot program and clarify parameters of service design and costs - Require states to collect data on district expenses and percentage of Title I funds - Provide states with technical assistance and guidance

Report	Problems	Recommendations
<p>Bertoni, D. (2007). <i>Highlights of a forum: Modernizing federal disability policy.</i> (GAO 07-934SP).</p>	<p>No federal system for coordinating disability programs.</p> <ul style="list-style-type: none"> - No lifetime picture of individuals with disabilities' needs. - More coordination and leadership of disability programs needed. - No private sector partnerships or incentives. - Disability populations and definitions vary - Data collection needs to be improved. - Multiple indicators are needed 	<ul style="list-style-type: none"> - Develop a coordinated government disability policy - Establish public-private partnerships and incentives - Improve the success of people with disabilities via higher education programs - Improve data collection for people with disabilities - Develop multiple indicators to measure success
<p>Nilsen, S.R., (2007). <i>Workforce Investment Act: Additional actions would further improve the Workforce System.</i> (GAO 07-1051T).</p>	<ul style="list-style-type: none"> -Comprehensive performance measures not established. - Data do not include all customer information - Data are not comparable across states and localities. -No information exists on what works and for whom - Impact evaluation of services was not funded until 2008 	<ul style="list-style-type: none"> - Continue improving the quality of performance data including WISPR - Align funding formulas labor market conditions and other factors (p. 12) - Create a systematic approach for adjusting performance levels -Conduct program evaluations and release findings in a timely way.
<p>Ashby, C.M. (2008). <i>Disconnected Youth: Federal action could address some of the challenges faced by local programs that reconnect youth to education and employment.</i> (GAO 08-313).</p>	<p>Program challenges included complex circumstances of participants, service gaps, funding constraints, federal grant management, 1 year timeframe to meet performance goals.</p>	<ul style="list-style-type: none"> -Labor should better ensure that states and WIBs have information and guidance. -Longer term contracts are intended.

Report	Problems	Recommendations
Scott, G.A. (2009). <i>Workforce Investment Act: Labor has made progress in addressing areas of concern, but more focus is needed on understanding what works and what doesn't.</i> (GAO 09-396T)	<ul style="list-style-type: none"> -Funding is not consistent with demands for services and available state funds -Only one impact evaluation of WIA programs has been completed -Evaluations of new grant initiatives will be problematic due to data and methodology issues 	No recommendations at this time.

Table 1 documents that there have been numerous problems with the program performance of federally-funded programs serving youth and youth with disabilities over the course of the last six or seven years. Recommendations to address the problems include conducting program evaluations, improving data collection, implementing corrective measures, coordinating efforts, increasing technical assistance and guidance, reconciling differences between program offices, and so on—leading the federal government to focus more intently on performance measures and take some drastic measures. For example, the President’s Budget Proposal for Fiscal Year 2008 noted that 141 non-security discretionary programs, worth \$12.0 billion, have been terminated after reviews to determine whether they were “among the Nation’s top priorities” and “effective and producing the intended results” (Executive Office of the President of the United States, Office of Management and Budget, 2007, p. 5). It was not clear, however, whether the problem was with the program intervention or with the program evaluation.

Congress recognized that federal grant management processes varied considerably across agencies and that “some Federal administrative requirements may be duplicative, burdensome, or conflicting.” (*P.L. 106-107*, §2). It passed *Public Law 106-107, the Federal Financial Assistance Management Improvement Act of 1999*, in order to “improve the effectiveness and performance of Federal financial assistance programs” and ultimately “the delivery of services to the public” (*P.L. 106-107*, §3). Impacted agencies (of which there are 26) were given eight years to develop a plan to streamline and simplify administrative and reporting requirements via electronic grant applications, reporting, and interagency coordination, and six years to evaluate the effectiveness of the Act. The Act “sun-setted” in November 2007, and although there have been some successes such as the Grant.gov web site which allows applicants to locate and apply for federal grants, other goals of the law have not been met—including the evaluation of the Act (Czerwinski, 2006; Posner & James, 2005).

Private Sector Calls for Improved Program Performance and Evaluation

In the private sector, foundations that were historically resistant to evaluating their funded projects changed their perspective on evaluating program effectiveness as public scrutiny of their activities increased, the field of evaluation evolved, and evaluation methodologies and expertise grew (Orosz, 2001). For over a hundred years, foundations in the United States did not evaluate their programs (Orosz, 2001). When evaluations became common in the 1980s, “expert” evaluators were hired by the foundations to evaluate grantees using a strict scientific methodology and to report back to the foundations with the findings (Orosz, 2001. p 4). This arrangement resulted in distrust of the evaluators (“spies”) by the grantees, problems applying scientifically rigorous

evaluation methodologies to programs “in the world of human development,” and the withholding of information and observations from the grantees until the end of the evaluation (Orosz, 2001, pp 5-6).

In the 1990s, participatory evaluation was born, in which the grantee, rather than the foundation, hired the evaluator, who worked with all grant stakeholders in formative (continuous improvement) and summative (outcomes) methodologies (Orosz, 2001). Formative evaluation provides information during the program to help improve it while summative evaluation is done after the program (or program cycle) has ended to determine if the program was effective and outcomes were achieved (Weiss, 1998a). Participatory evaluation results are not as definitive as a strictly scientific approach, and some objectivity may be lost by the evaluator, but grantees are more involved in the evaluation process, and the program may be improved as it progresses (Orosz, pp. 7-8). However, participatory evaluation does not necessarily mean that grantees are comfortable with the evaluation process or the evaluator (based on the researcher’s personal experiences).

Foundations today are looking at improving the evaluation process as well as grant outcomes and results. In March 2007, the Hauser Center for Nonprofit Organizations at Harvard University and the Nonprofit Finance Fund convened a discussion of 50 nonprofit leaders on the topic of finance. The result was “support for a fundamental shift in the way funders support the nonprofit sector” (Buechel, Keating & Miller, 2007, p. 5). In the survey that was distributed before the meeting, 82 respondents agreed on a number of changes that need to be implemented in order to improve the efficiency and impact of nonprofits, including these relating to evaluation:

1. Grantees and foundations should jointly decide evaluation tools and metrics (p. 6)
2. Standardization of grant applications and reporting *across funding organizations* is desirable [emphasis added] (p. 6)

Other findings from the meeting included:

3. Grant applications and reporting need to be simplified (p. 5)
4. Performance management, which now is “focus[ed] on current activities and outputs, compliance oriented, command and control oriented, [with] crude, simplistic measures” should change to “root-cause analysis, outcome/results orient[ation], performance-driven management, [and a] totally quality management philosophy” (p. 12).
5. Accountability should move away from “reporting driven by regulations and funders” to “active, informed board reporting to key constituents” (p. 12)

A survey of philanthropies by *The Chronicle of Philanthropy* survey found “increased introspection” (Barton & Wilhelm, 2007, p. 5) among some respondents, including conducting internal reviews “to do a better job of solving emerging problems” (p. 2) and “being pretty self-critical” about mistakes made, lessons learned, communicating with the public, and other concerns. Other philanthropies reported focusing their giving more tightly, such as on disaster-related or diversity-related giving.

A number of current foundation practices relate to evaluation. The W.K. Kellogg Foundation has its own evaluation philosophy, expectations, handbook (1998), and logic model development guide (2001) for use by grantees. The United Way of America provides a commonly used logic model and other guidelines in *Measuring Program*

Outcomes: A Practical Approach (Hatry, van Houton, Plantz & Taylor, 1996); its *Standards of Excellence* (2005) include “impact strategies” such as “measuring, evaluating, and communicating results” that “may cause United Way and its partners to re-think, change, or adjust strategies, actions, and investments” (p. 11). The United Way is now moving to a new framework, “Results Based Accountability” from the Fiscal Policy Studies Institute (Friedman, 2005) which presents a different evaluation framework and new terminology in practical, understandable language.

Sometimes foundation accountability requirements move beyond documentation of progress toward goals. In February 2002, the Grable, Heinz, and Pittsburgh foundations created a sensation when they very publicly (but temporarily) suspended over \$3.5 million to Pittsburgh public schools in order to “ensure that our investments in the school system realize our shared goals of enhancing student achievement and promoting fiscal discipline” citing “a sharp decline of governance, leadership and fiscal discipline” in the school system (Lee & Elizabeth, 2002, 6th paragraph).

Funding Programs and Services for Transition-Age Youth

Financing programs and services for youth of transition age (14-25) is becoming more challenging for non-profit programs seeking federal, state, and private funding in the United States (Chambers, 2006). For example, the FY 2008 Budget for education, training, and employment reflected White House priorities of supporting the war in Iraq, balancing the budget, and “budget reforms that will improve transparency and accountability in Government spending” (Executive Office of the President of the United States, Office of Management and Budget, 2007, p. 5). The 2008 budget Overview promised “significant new resources and other reforms” for education including more

funding for high schools, Pell Grants, and Academic Competitiveness Grants (p. 4). However, Budget Table 9.9 (Composition of Outlays for the Conduct of Education and Training) indicated that proposed FY 2008 federal expenditures for education, training, and employment dropped to 2.8% of total federal outlays or \$61.7 billion in constant (2000) dollars from a high of 4.4% or \$94.4 billion in FY 2006 (p. 191).

Since 2007, the United States has entered a severe recession. The 2010 budget for the Departments of Education and Labor show slight increases (approximately 3%) over the 2009 budget (U.S. Office of Management and Budget, n.d.b, n.d.c) but financial stresses such as decreased state education funds and increased Unemployment Insurance claims will reduce their impact.

Other analyses of the proposed federal budget indicate lower funding for initiatives that serve youth. The Center for Budget and Policy Priorities found that, in addition to the cuts identified in the FY 2008 federal budget, an additional \$114 billion are planned to be cut from domestic discretionary programs in 2009 through 2012 including elementary and secondary education, job training, child care and dozens of other programs (Greenstein, 2007, p. 5). An Urban Institute report that tracked federal spending on children (under the age of 19) from 1960 projected through 2017 concluded that “the analysis of historical and future trends in the federal budget reveals that children are a diminishing national priority” (Carasso, Steuerle & Reynolds, 2007, p. 3).

An analysis of the FY 2007 budget’s impact on the nonprofit sector from The Aspen Institute found that “the budget proposals... recently submitted to Congress, covering fiscal year 2007 and beyond, would put new demands on the nation’s private, nonprofit organizations... [and] reduce the federal support going to these organizations to

provide services” (Abramson, Salamon, & Russell, 2006, p. 1). Further, it found that budget cuts “would affect nearly every program area of interest to nonprofits” (p. 2) with the biggest cuts going to the area of education and research (p. 4).

However, *The Chronicle of Philanthropy* reports that a new study to be released soon by The Aspen Institute found “that in its early stages of making spending decisions, Congress approved an increase of three times more money for programs of interest to charities than President Bush sought in the budget request he sent to Capitol Hill in February” (Schwinn, 2007, third paragraph). Most of the Congressional increases will go to health care programs and income security programs. Other programs such as the arts, education, nutrition, and scientific research would be cut, but not as much as in the Bush plan. *The Chronicle* also reports that another Aspen Institute report found that more charities are qualifying for Medicaid funds in order to serve their clients, even if their primary missions are not directly related to health (Schwinn, 2007), despite Medicaid’s difficult qualifying process and uncertain funding availability.

A strained federal budget will mean fewer programs and fewer evaluations. Youth organizations will want to put as many resources as possible into direct program services. Federal budget cuts also have a trickle down effect on the budgets of state and localities.

At the state level, before the country went into the current recession, budget and expenditure trends were mixed (McNichol, Johnson, & Farkas, 2007; National Conference of State Legislatures, 2007) or negative (Czerwinski & McCool, 2007; Head, Sigritz, & Mazer, 2007). The Center on Budget and Policy Priorities predicted that growth would slow for both state revenues and expenditures which could “foreshadow

problems ahead for state budgets” (McNichol, Johnson, & Farkas, 2007, p. 3) as did a joint report by the National Governors Association (NGA) and the National Association of State Budget Officers (NASBO) (Head, Sigritz, & Mazer, 2007). Reports of “fiscal stress” (Czerwinski & McCool, 2007, p. 4) were surfacing in California, Illinois, Maryland, and New Jersey due to projected budget shortfalls and overdue payments into pension funds. NGA and NASBO projected decreases in budget stabilization funds and significant challenges in funding and providing health care for the upcoming fiscal year (Head, Sigritz, & Mazer, 2007).

The National Conference of State Legislatures (NCSL) had similar concerns and quoted U.S. Comptroller General David Walker who described “a grim outlook for America’s fiscal future” due to the shifting of billions of dollars of Medicare, Medicaid, Social Security and other costs to the states over the last three years, a trend which is expected to continue (NCSL, 2006). The following year, NCSL reported that although “revenues eclipse estimates in most states,” “sales and use tax collections are causing some concerns” (2007, p.1) and a lobbying priority will be to “curb unfunded federal mandates, which strain state budgets by billions of dollars” (2007, p.2). The Government Accountability Office projected that, without policy changes, state expenditures would exceed state revenues in the next decade primarily due to burgeoning health care costs, specifically Medicaid and state employee health insurance (Czerwinski & McCool, 2007).

The latest *Fiscal Survey of States* (Husch, Cheung, Cummings, Mazer, & Sigritz, 2009) conducted by the National Governors Association and the National Association of State Budget Officers states that “the 50 states are facing one of the worst fiscal periods

in decades.” (Executive Summary). Dropping tax revenues and rising job losses have caused budget deficits and increased demands for social services. Stimulus funds and budget cuts will allow some states to hold the line, but budget cuts and weak fiscal conditions are expected to continue at least into 2010 or beyond.

Before the recession, the funding scenario for corporate, community, and independent foundations was more upbeat. A report by the Foundation Center found that the United States was “experiencing a new ‘golden age of philanthropy’”—despite “the less predictable current environment” of potential terrorist attacks, natural disasters, and stock market volatility—because of the increasing numbers of private and community foundations, growing foundation assets due to the stock market, and “high profile donations of vaccines and medicines by pharmaceutical companies” (Lawrence, Austin & Mukai, 2007, p. 2). Growth in foundation giving increased by 14.3 percent in 2005 and 11.7 percent in 2006, with similar growth projected for 2007 (pp. 1-2). However, foundations and charities are also requiring stronger evaluations as part of their funding, and it is not clear how what portion of their funds will be allocated to organizations serving youth ages 14-25.

The most notable donation by far has been Warren Buffet’s commitment to donate over \$40 billion to five foundations with the lion’s share going to the Bill and Melinda Gates Foundation (Loomis, 2006). In addition, 112 of the nation’s wealthiest foundations surveyed by *The Chronicle of Philanthropy* increased their assets, diversified their stock portfolios, and planned to increase their grant awards significantly in 2007 (Barton & Wilhelm, 2007).

However, the Council on Foundations found that the 127 members who responded to its February 2009 survey reported a 28% decline in assets over the previous year. Members also reported changes to asset allocations, investment consultants, investment managers, and attempts to reduce portfolio investment fees. *EducationWeek* also reported declines in educational endowments with several philanthropies planning to decrease grant awards this year (Robelen, 2009). Some foundations will maintain planned funding levels, while others, such as the Gates Foundation, plan to give more this year. Nonprofits and private schools are anticipating—or experiencing—dropping revenues based on dwindling assets and declining fundraising (Robelen, 2009).

A new survey report from Giving USA Foundation (2009, June 10) found that donations to charities in 2008 dropped 2% over 2007, which was only the second time since 1956 that giving declined. Human services charities reported that 54% experienced an increase in need, 60% were cutting expenses and/or staff, and organizations serving children and youth and those providing food shelter, and clothing were underfunded or severely underfunded. Meanwhile the Nonprofit Finance Fund was circulating recommendations for surviving the recession based on lessons learned in 2001 (Nonprofit Finance Fund, 2008, February 13).

Summary

Both the federal and private sectors are requiring better returns on their youth program investments as evidenced by an increased focus on positive performance and results. Meanwhile, federal funding continues to tighten due to the current economic recession, state budgets are stretched to the limit (or beyond), and the assets of many philanthropies are losing value. Nonprofits and other youth service providers are

experiencing difficulties in fundraising and anticipate strong competition for available funds.

The increased demand for results and the decrease in available funds create stress for youth service providers. The importance of “positive” evaluations and their potential impact on future funding are concerns that will influence evaluation planning and implementation. It is likely that these strains will influence procedural and interpersonal aspects of the working relationship between evaluators and their youth service provider clients. The current climate for evaluations of youth programs is one of the primary reasons for this study.

Performance of Programs and Initiatives Serving Transition-Age Youth

The stakes are high for providers of youth services as funders, both governmental and private, are taking a hard look at how effectively their funding is used (The U.S. Office of Management and Budget, n.d.a; Buechel, Keating & Miller, 2007) and how to more effectively address “difficult social problems” (Buechel, Keating, & Miller, 2007, p. 5).

The track record of programs and initiatives serving transition-age youth in the United States is not impressive across a wide spectrum of service providers ranging from large organizations (postsecondary institutions, public schools, the workforce development system) funded by tuition, fees for services, and federal/state government budgets, to small, single-site programs operated by community and faith-based organizations funded by a patchwork quilt of grants and donations. Since average income and other indicators of well-being have been linked to levels of education completed (Kirsch, Braun, Yamamoto, & Sum, 2007; Laird, Kienzi, DeBell & Chapman, 2007;

Swanson, 2007; Olson, 2007; Carey, 2004; Gladieux & Swail, 1998) and skills attained (Economic and Development Review Committee, 2007; Kirsch, Braun, Yamamoto, & Sum, 2007; Partnership for 21st Century Skills, n.d.; Gandal, Slover, Kraman & Pache, 2004), ensuring that young people successfully transition from education to employment and the adult world is critically important, not only for themselves, but for the future of the United States (Economic and Development Review Committee, 2007; Kirsch, Braun, Yamamoto, & Sum, 2007; New Commission on the Skills of the American Workforce, 2007; National Center for Public Policy and Higher Education, 2006; Gladieux & Swail, 1998; Partnership for 21st Century Skills, n.d).

The following sections will examine the performance of systems serving transition age youth including (a) education, the primary system serving this age group, composed of institutions of higher education and public/private K-12 schools, (b) second chance systems such as the Workforce and Adult Basic Education/General Educational Development systems, and (c) smaller, discretionary programs administered by community and faith-based organizations, nonprofits, and other organizations.

Performance of Higher Education

The Organisation [sic] for Economic Co-Operation and Development found that “the United States system of higher education is widely seen to be the best in the world” because its

colleges and universities offer more choice, their graduates receive greater wage premiums, and they attract more than twice as many foreign students as any other country....However... the [Spellings] Commission and others have pointed to serious problems in the areas of accountability, quality, transparency, cost control, diversity, and many other aspects of higher education (Economic and Development Review Committee, 2007, p. 126).

Using data from the United States Census Bureau, the U.S. Department of Education, and other sources, a number of organizations have concluded that colleges and universities in the United States can do better in improving graduation rates and other educational outcomes of students (New Commission on the Skills of the American Workforce, 2007; Casner-Lotto, 2006; Horn, 2006; National Center for Public Policy and Higher Education, 2006; Barton, 2005; Carey, 2004; Gladieux & Swail, 1998). For example, The National Center for Public Policy and Higher Education (2006) graded all 50 states on six parameters related to higher education:

1. Preparation for higher education – including high school completion, course taking, student achievement, and teacher quality: Seven states (14%) received an “A” (p. 18)
2. Participation in higher education – including access by age, race/ethnicity, family income: 17 states (34%) received an “A” (p. 18)
3. Affordability: No state received an “A” or “B” (p. 18)
4. Completion – Although completion rates are slowly improving, one third of two and four year entering freshman do not complete degree requirements: 14 states (28%) received an “A” (p. 198)
5. Benefits – “Educational capital” of states as measured by percentage of residents with a bachelors degree or higher (p. 14): 17 states (34%) received an “A” (p. 18)
6. Learning – Because comparable data were not available across states to measure literacy levels, “graduates ready for advanced practices,” and performance of college graduates, “incompletes” were awarded to all states

except nine (18%) that received “pluses” for participating in a national demonstration project on college learning or the State Assessment of Adult Literacy (SAAL) (p. 23)

Evaluating institutions of higher education is difficult for a number of reasons. College graduation rates vary significantly based on factors such as type of institution (ACT, 2007), institutional selectivity (ACT, 2007; Horn, 2006), and family income, gender, and race/ethnicity (Horn, 2006; Carey, 2004; Gladieux & Swail, 1998). Difficulties were noted in comparing data across institutions for outcome measures other than graduation rates (Horn, 2006) because of data collection or availability issues (National Center for Public Policy and Higher Education, 2006; Gladieux & Swail, 1998).

Performance of K-12 Education Systems

The public and private K-12 education systems in the United States serve approximately 55 million students a year (National Center for Education Statistics, 2007). The Organisation for Economic Co-Operation and Development summarizes schooling in the United States as follows:

Primary and secondary education outcomes are unsatisfactory. A country’s ability to compete in an ever more integrated world economy crucially depends on a highly educated workforce. It is thus a matter of concern that US students are outperformed in international tests by their peers in many other countries. It is also a concern that many students seem underprepared for work and higher education. Although large achievement gaps persist between population groups, performance is broadly unsatisfactory, including among affluent and academically successful students, and hence appears to reflect above all system-level weaknesses. Available evidence does not unambiguously establish the reasons for this. (Economic and Development Review Committee, 2007, p. 14).

Determining the performance of public K-12 schools is difficult due to problems in data collection such as quality, availability, and transparency of data across states

(Chudowsky, Chudowsky & Kober, 2007; Robertson, 2007; Bellis, 2005; Curran, 2005; NISS/ESSI Task Force on Graduation, Completion, and Dropout Indicators, 2005; Shaul, 2005; Shaul, Edward, & Fucile, 2004; Sum, Harrington, Bartishevich et al, 2003), lack of uniformity in formulas for calculating key indicators such as dropout and graduation rates across states (NISS/ESSI Task Force on Graduation, Completion, and Dropout Indicators, 2005; Shaul, 2005), and “perverse incentives” such as using student transfers to manipulate graduation indicators (NISS/ESSI Task for on Graduation, Completion, and Dropout Indicators, 2005, p. 18; Shaul, 2005, p. 25; Hall, 2005).

Despite these issues, the National Center for Education Statistics’ most recent calculation of the national freshman graduation rate (the percent of freshman who graduate with a regular diploma four years after starting the 9th grade) for 2006 was 73.2% (Planty, Kena, & Hannes, 2009). Fifteen states had rates of 80% or above, and 12 states had rates less than 70% (Planty, Kena, & Hannes, 2009, pp. 12-13). National event dropout rates describe the percentage of students who dropped out of school between the beginning of one school year and the beginning of the next school year. The 2006-2007 overall event dropout rate was 9% (Planty, Kena, & Hannes, 2009, pp. 14-15).

Other organizations have taken issue with the NCES figures and have placed the national graduation rate at 66.1% to 71.0% (Barton, 2005, Shaul, 2005; Hall, 2005). Some researchers have been more pointed in their concerns about the calculation and reporting of graduation rates in regards to racial and gender disparities, or as the Civil Right Project at Harvard University put it, “How has the misleading and incomplete reporting of this issue obscured both the magnitude and racial dimensions of the crisis?” (Orfield, Losen, Wald, & Swanson, 2004, p. 7)

The data cited above shows that, despite significant resources, evaluating higher education and K-12 schools is difficult due to problems with accuracy and comparability of data. Further, the data that is available appears to indicate that the two systems are not doing a good job in producing graduates who are adequately prepared for employment and further education. Calculating graduation rates for students with disabilities is an additional challenge since the National Center for Education Statistics and the Office of Special Education Programs use different formulas for calculating graduation rates (National Dropout Prevention Center for Students with Disabilities (NDPCSD), 2006). These difficulties support the need for more consistency in conducting and reviewing program evaluations.

For example, states are required to prepare State Performance Plans that contain indicators showing that they are meeting the goals and are in compliance with the Individuals with Disabilities Education Act. Indicator 1 for Part B of the Act (for children ages 3 to 21 receiving special education and related services) requires each state or jurisdiction to calculate the “percentage of youth with IEPs graduating from high school with a regular diploma compared to the percent of all youth in the State graduating with a regular diploma” (NDPCSD, 2006, p. 1). In an analysis of IDEA State Performance Plans submitted in December 2005, the National Dropout Prevention Center for Students with Disabilities found a number of discrepancies in the way graduation data was calculated for Indicator 1 by the 60 jurisdictions that are subject to IDEA (including the 50 states, the District of Columbia, the Bureau of Indian Affairs, and U.S. territories and commonwealths) (NDPCSD, 2006). Some jurisdictions used a formula recommended by the Office of Special Education Programs in the U.S. Department of Education, others

used a National Center for Education Statistics formula, and others used a combination of the two. NDPCSD also found between-state variations in definitions of graduation (such as variations in diploma options for students with disabilities) and 19 states or jurisdictions that were unable to report data for the 2004-2005 school year. (Some states reported data for school years 2002-2003 or 2003-2004, or no data at all.) Based on the data submitted, state graduation rates for students receiving special education services ranged from 4% to 92.5% (NDPCSD, 2006, p. 3).

The U.S. Department of Education Determination Letters on State Implementation of the IDEA released in June 2007 found continuing problems with the calculation of graduation rates for students with disabilities (Indicators 1 and 2, Part B, Office of Special Education Programs, 2007). OSEP (2007) also found problems with the calculation of performance indicators and baseline data such as rates of suspension and expulsion (Indicator 4, Part B), numbers of students with Individual Education Programs containing “transition services that will reasonably enable the student to meet the post-secondary goals” (Indicator 13, Part B), and post-school outcomes for students with disabilities who have left school (Indicator 14, Part B).

The National Governors Association, The National Center for Education Statistics, and the Institute of Education Sciences, among others, are working with the states to create state database systems that will allow calculation of student progress data that will be comparable across states and educational programs (Curran, 2005; Hicks, 2004, Institute of Education Sciences, 2007). How long this process will take is unclear. What is clear is that it is difficult or impossible to evaluate the performance of regular

and special education without accurate outcome indicators, despite the pressure on the system to produce them.

A number of organizations, in addition to OECD, have concluded that K-12 students who graduate from high school are unprepared for life after school. Achieve, Inc. reports that students earn a high school diploma “without mastering the knowledge and skills they need to succeed after graduation” and cites data that estimate that only “32% of high school graduates, 20% of African Americans, and 16% of Latinos are prepared to succeed in college” (Gandal, Slover, Kraman & Pache, 2004, p. 5). The Partnership for 21st Century Skills (2004) states that “there is a profound gap between the knowledge and skills most students learn in school and the knowledge and skills they need in typical 21st century communities and workplaces.”

The New Commission on the Skills of the American Workforce found that “the core problem is that our education and training systems were built for another era, an era in which most workers needed only a rudimentary education.... We can get where we must go only by changing the system itself” (New Commission on the Skills of the American Workforce, 2007, p. 8). A survey of over 400 employers nationwide conducted by The Conference Board, Corporate Voices for Working Families, The Partnership for 21st Century Skills, and The Society for Human Resource Management, found that high school graduates (along with 2 and 4 year college graduates) to be “woefully unprepared for the demands of today’s (and tomorrow’s) workplace” (Casner-Lotto, 2006, p. 9).

Accurate and appropriate evaluations of college and public school programs and services would be helpful in identifying specific parameters that need improvement as well as programs and services that work well and could be replicated across the country.

Performance of “Second Chance” Programs

In order to improve outcomes for transitioning youth and adults, a number of supplemental programs and “second chance” options have emerged in the United States to move people lacking the necessary skills and knowledge to economic self-sufficiency. One of the largest second chance programs is the national workforce development system funded by the Workforce Investment Act. WIA serves low-income youth and unemployed or low-income adults in order to “build the talent necessary for individuals, companies, and regions to succeed in the 21st Century economy” (Employment and Training Administration, 2007, inside cover).

The two WIA programs most likely to be accessed by youth ages 14-25 are the Youth Program for low-income youth ages 14-21 with barriers to employment and the Adult Program for adults ages 18 and older who are unemployed, underemployed, or are members of at-risk groups. Results for program year 2006 data (through March 31, 2007) were mixed (Employment and Training Administration, 2007, p. 7):

1. 58% of exiting youth were placed in employment or enrolled in postsecondary education or training which was below the Government Performance and Results Act (GPRA) goal of 60%.
2. 40% of exiting youth attained a high school diploma, GED, or educational certificate which met the GPRA goal of 40%.
3. 71% of exiting adults entered employment which was below the GPRA goal of 76%.
4. 83% of exiting adults retained employment for at least two quarters following their exit quarter which exceeded the GPRA goal of 82%.

Another “second chance” system is composed of Adult Basic Education/General Educational Development (ABE/GED) programs operated by public school systems, community colleges, and others. The GED test battery offers people who were not able to complete high school the opportunity to earn a high school equivalency credential. The American Council on Education (ACE) and its offspring, the General Educational Development Testing Service (GEDTS), reported that nearly 68% of the 419,000 people who completed the battery of GED tests in 2006 passed with “scores equal to or above those earned by 40% of graduating high school seniors” while cautioning that policies on prescreening and retesting vary across GED programs and can affect passing rates (ACE & GEDTS, 2007, p. ix). 72% of completers passed the test battery in 2005 (ACE & GEDTS, 2007, p. xi).

A number of sources do not consider the GED to be equal to a high school diploma, although they agree it is better than being a dropout (Arenson, 2004; Tyler, 2004; Education Portal 2003-2007; American Youth Policy Forum, 1998). GED recipients had better employment rates and earnings than dropouts but had lower employment rates, lower earnings, and higher turnover rates than high school graduates. They were also less likely to complete a 2- or 4- year college program than high school graduates (AYPF, 1998). There is now a growing concern that GED programs may encourage students to drop out of high school (Tyler, 2004). A rigorous evaluation of the ABE/GED system would be helpful in determining areas of success and needed improvements.

Performance of Discretionary Grant-Funded Organizations

As described above, the huge systems of higher education and public K-12 with their many resources have great difficulty providing accurate data on outcomes and performance. It should come as no surprise, therefore, that many youth programs funded by discretionary grants, often to nonprofits and community-based organizations, have difficulty in demonstrating their effectiveness, particularly since many use data supplied by the larger systems for comparison purposes. For example, a What Works Clearinghouse review (2007) of school- and community-based dropout prevention programs showed less than encouraging results.

WWC reviewed 58 studies of 16 programs and found five programs that did not meet WWC evidence requirements due to lack of data or an acceptable study methodology. The WWC report did not describe the methodologies used by the programs but stated that only randomized control trials or quasi-experimental designs “with evidence of equating between the treatment and comparison groups” were eligible for WWC review (WWC, 2007, p. 10). Of the remaining 11 programs, seven studies met WWC evidence standards “without reservations,” and nine studies met WWC standards with reservations (2007, p.10). The 11 programs were then rated for program effectiveness in one or more of three domains. WWC effectiveness ratings (2007, p. 2) for the Completing School domain indicated that one program (of eight) had “potentially positive effects” while seven had “no discernable effects.” Ratings in the other domains were better, with one program (of seven) showing positive effects and four showing potentially positive effects for Staying in School, and five programs (of nine) showing potentially positive effects in Progressing in School.

Regardless of the methodology used, published evaluations of privately funded programs serving youth 14-25 may find positive effects, promising practices, or lessons learned but usually cannot establish causality or state unequivocally which interventions are effective (WWC, 2007; Branch-Smith, Gray, Fruchter, Hernandez, Joselowsky, Nichols-Solomon, Simmons, & Warren, 2006; Bain, 2003; Blank, Melaville, & Shah, 2003). For example, the Coalition for Community Schools' review of evaluations of 20 local, state, and national community school initiatives found positive effects on student learning, family engagement, school effectiveness, and community vitality (Blank, Melaville, & Shah, 2003) but noted problems with lack of comparability across programs due to differences in methodologies and lack of model fidelity in implementing program models.

Some of the most difficult interventions to evaluate are initiatives funded to improve outcomes for youth by changing the systems that serve them. The Schools for a New Society Initiative, for example, was funded by the Bill and Melinda Gates Foundation and Carnegie Corporation of New York to reform the institution of the high school "to ensure success for all students" (Branch-Smith, Gray, Fruchter, Hernandez, Joselowsky, Nichols-Solomon, Simmons, & Warren, 2006, p. 69) After five years of implementation in seven cities, the Technical Support Team found that "great progress" was being made toward the vision, including a framework for effective redesign and a number of "emerging lessons", but that "there [was] certainly work still to be done" (p. 70).

The Annie E. Casey Foundation funded "an ambitious social experiment" called New Futures to "measurably improve school achievement, reduce adolescent pregnancy

and school dropout rates, and increase young people's gainful employment after high school" in five cities over five years (Center for the Study of Social Policy, 1995, p. i). The "single core strategy [was] broad institutional change in the way services and supports are provided to at-risk youth as determined by a newly constituted local governing body called a collaborative" (CSSP, 1995, p. i). This "experiment using private money to leverage public policy and public financing on a major social issue of the day" (CSSP, 1995, p. ii) was ultimately unsuccessful in improving dropout rates, failure rates, the achievement gap between white and black students, and youth pregnancy and employment rates. However, important lessons were learned about the challenges and processes of cross-systems change—among them, the dangers of "hurried implementation of weak plans that lacked high level endorsements" (CSSP, 1995, p. 184)—lessons that can be applied to design, implementation, and analysis of evaluation themselves.

The federal government also funds competitive, discretionary programs such as the Comprehensive School Reform (CSR) initiative. *Comprehensive School Reform and Student Achievement: A Meta-Analysis* (Bain, 2003) found that "the average student from a CSR school outperformed about 55% of the children who did not attend a CSR school" (p. 1). However, the analysis identified "two major biases in the research literature" (p.1) and noted "that the ideal conditions for the unbiased studies of a truly experimental design are much more difficult and costlier to maintain especially in the area of education" (p.1)—additional evidence that conducting evaluations in complex systems is difficult.

Recently published studies of programs funded by appropriated and discretionary funds continue the trend of no effects or mixed results. *Education Week* (Viadero, 2009) reported no effects for four studies that included an impact evaluation of a federal student mentoring program, a comparison of elementary math curricula, an evaluation of ten commercial software products, and a comparison of student achievement for students whose teachers were trained with traditional or alternative programs.

Summary

The performance of K-12, postsecondary, and workforce development systems serving youth ages 14-25 is characterized by difficulties collecting accurate, appropriate data and achieving planned outcomes such as employment and high school or college graduation rates. Youth programs funded through discretionary grants experience similar difficulties. Youth service providers will be anxious to show positive results for their programs, and this anxiety will influence procedural and interpersonal aspects of the working relationship between evaluators and their youth service provider clients—a additional reason for this study.

The Challenges of Youth Program Evaluation

Determining the effectiveness of a program or intervention may be difficult because the evaluation methodology or data collection is problematic--as noted above in evaluations of institutions of higher education, the public schools, systems change initiatives, and local or more targeted programs. Evaluators have identified a number of challenges to implementing effective evaluations such as

1. rigid organizational procedures, settings, or leadership (Taut & Alkin, 2003; Weiss, 1998a, Cecil, 2004),

2. problems with data collection, management, analysis, or reporting
(Donaldson, Gooler & Scriven, 2002; Taut & Alkin, 2003; Stoecker, 2007;
Kegeles, Rebchook, & Tebbetts, 2005; Mulroy & Lauber, 2004),
3. limited methodological capacity of staff (Hassenfeld & English, 1987;
Stoecker, 2007; Kegeles, Rebchook & Tebbetts, 2005, Mulroy & Lauber,
2004);
4. disagreement about who should conduct the evaluation (Stoecker, 2007),
5. lack of cooperation, collaboration or communication among staff (Weiss,
1998a: Donaldson, Gooler, & Scriven, 2002: Hasenfeld, & English, 1974;
Gilliam, Davis, Barrington, Lacson, Uhl, & Phoenix, 2002),
6. Lack of resources such as funding and staff time (Gilliam, Davis, Barrington,
Lacson, Uhl, & Phoenix, 2002; Hasenfeld, & English, 1974; Taut & Alkin,
2003; Stoecker, 2007; Cecil, 2004; Mulroy & Lauber, 2004)
7. unexpected/unanticipated evaluation results or consequences(Donaldson,
Gooler, & Scriven, 2002; Taut & Alkin, 2003)
8. conflicts with funders ((Stoecker, 2007; Gilliam, Davis, Barrington, Lacson,
Uhl, & Phoenix, 2002)
9. philosophical misapplication of theories such as rationalism, constructivism,
or post-modernism (Weiss, 1998a; Thayer, 2006)
10. evaluation results that were not used (Patton, 2008; Donaldson, Gooler, &
Scriven, 2002)

It should be no surprise, given the challenges identified, that youth programs have experienced challenges with evaluation implementation. The Coalition for Community

Schools' review of evaluations of community school initiatives, for example, included a candid discussion of "technical limitations" and challenges in implementing program evaluations:

All too often, funders expect sophisticated outcome evaluations, but overlook the resources and capacity needed to conduct them. Few programs have the capacity to track individual outcomes...Clearly more resources need to be invested in community school research both to refine evaluation methodologies and to enable a better understanding of the factors at play in high-quality community schools. Only when we better understand these intricacies will we be able to target efforts to expand and sustain their effects. (Blank, Melaville, & Shah, 2003, p.46)

Technical limitations for the Coalition's review included variations in school evaluation methodologies such as reviews of program records, participant surveys, and comparisons of results from "selected" community schools to those from similar non-community schools (p. 46). Problems in implementing the various methodologies were also noted such as comparison schools becoming community schools during the course of the evaluation (p. 47), "sites that differed significantly in the duration, quality, and kind of activities they provided" (p. 47), and difficulties collecting parental consent forms for student participation (p.47).

Bain's meta-analysis of *Comprehensive School Reform* initiatives also found problems with determining effects and concluded that

CSR models [are] an effective way of improving student achievement. The research base continues to evolve, and sustained research efforts need to be applied to improve quality of research on CSR effects and to examine these effects as closely related to CSR implementation and specific school contexts (2003, p.2).

The Government Accountability Office (GAO) is the "audit, evaluation and investigative arm of Congress" and evaluates federal programs and policies as one of its primary responsibilities (Nilson, 2007, p. 25). As noted previously, a number of GAO

reports and statements of Congressional testimony indicate ongoing problems with effective program evaluation, including youth-serving initiatives and services funded by *the Individuals with Disabilities Education Act, No Child Left Behind Act, Workforce Investment Act*, Office of Juvenile Justice and Delinquency Preventions, and others. See Table 1 for a sampling of problems and recommendations identified by the GAO with these programs.

Ideological biases, opposing views on appropriate methodologies, and factional disagreements may further degrade the quality of program evaluations. In April 2007, *USA Today* reported that U.S. Secretary of Education Margaret Spellings presented “a \$100 million proposal to use taxpayer-funded vouchers to send public school students to private schools”—four days after a study commissioned by the Department of Education was released which found that “public school students often do better and sometimes a lot better than private-schoolers” (Toppo, 2007). Secretary Spellings dismissed the report, calling it “irrelevant”, “small, and flawed” (Toppo, 2007). The reporter concluded:

In the end, it was a pretty good metaphor for the state of educational research: More than five years after President Bush’s No Child Left Behind law told educators to rely on ‘scientifically based’ methods, the science produced is often inconclusive, politically charged or less than useful for classroom teachers. And when it is useful, it often is misused or ignored altogether (Toppo, 2007).

The U.S. Department of Education’s emphasis on “scientifically based evaluation methods” was formally announced in a “notice of proposed priority” published in the November 4, 2003, *Federal Register* (Paige, 2003, p. 62445). The notice created a firestorm among educational researchers and evaluators who took issue with the policy’s language describing acceptable experimental designs for evaluating program projects funded by the department as in the passage below:

Proposed evaluation strategies that use neither experimental designs with random assignment nor quasi-experimental designs using a matched comparison group nor regression discontinuity designs will not be considered responsive to the priority when sufficient numbers of participants are available to support these designs. Evaluation strategies that involve too small a number of participants to support group designs must be capable of demonstrating the causal effects of an intervention or program on those participants. (Paige, 2003, p. 62446)

Organizations as varied as the American Evaluation Association (Krueger, 2003), the American Educational Research Association (Levine, 2003), the National Education Association (Moody, 2003), and the Visitor Studies Association (Munley, 2003) objected to the extremely narrow range of methodologies considered “responsive” and the resulting emphasis on randomized control group trials. The American Evaluation Association’s letter and attached statement (Krueger, 2003, 1st paragraph of statement) emphasized that “[w]e would like to help avoid the political, ethical, and financial disaster that could well attend implementation of the proposed policy” and pointed out that AEA encourages “competent practice through our *Guiding Principles for Evaluators* (1994), *Standards for Program Evaluation* (1994), professional training, and annual conferences,” (Krueger, 2003, 1st paragraph of statement) that other experimental designs are capable of showing causation, particularly when the use of Randomized Control Trials (RCTs) is unethical or data is insufficient, and that “fettering evaluators with unnecessary and unreasonable constraints would deny information needed by policy-makers” (Krueger, 2003, Statement Item 2). Other organizations (Levine, 2003; Moody, 2003) suggested that the Department align its priority with the definition of “scientifically valid education evaluation” set forth in *Public Law 107-279, Education Sciences Reform Act of 2002*:

The term “scientifically valid education evaluation” means an evaluation that—

- (A) adheres to the highest possible standards of quality with respect to research design and statistical analysis;
- (B) provides an adequate description of the programs evaluated and, to the extent possible, examines the relationship between program implementation and program impacts;
- (C) provides an analysis of the results achieved by the program with respect to its projected effects;
- (D) employs experimental designs using random assignment, when feasible, and other research methodologies that allow for the strongest possible causal inferences when random assignment is not feasible; and
- (E) may study program implementation through a combination of scientifically valid and reliable methods (20 USC 9511).

In its final notice of proposed priority, the Department noted that “almost 300 parties submitted comments” (Paige, 2005, p. 3586) including 242 who said that “the choice of a research method must be determined by the goal or question being asked” (Paige, 2005, p. 3589). Over 170 respondents took issue with various aspects of the emphasis on random assignment of subjects (Paige, 2005, pp. 3588-3589). A few respondents, 29 and nine respectively, agreed with the Department’s position on randomized control trials and preference for external evaluators (Paige, 2005, p. 3588 and 3589). However, the final priority included no changes other than (a) clarifying that the priority would be used “only” for evaluations of competitively funded discretionary grants used to assess interventions in order to determine the impact on outcomes, (b) include a two stage review process for grant proposals, and (c) add definitions for technical terms (Paige, 2005, p. 3589).

Not surprisingly, therefore, the What Works Clearinghouse (WWC), funded by the U.S. Department of Education to identify effective evidence-based research (as described in *Public Law 107-110*, the reauthorization of the *Elementary and Secondary Education Act* commonly known as the *No Child Left Behind Act of 2001* and *Public Law*

107-279 Education Science Reform Act of 2002), has come under fire. *Education Week* reported that “critics call the What Works Clearinghouse the ‘nothing works’ clearinghouse” (Viadero, 2006, 2nd paragraph) and described complaints from researchers about the standards used to evaluate programs and timeliness of reporting findings. Alan Schoenfeld, a professor at the University of California, Berkeley, accused the Clearinghouse of “censorship” in its review of mathematics curricula and was “forced to conclude that the enterprise as currently constituted has neither intellectual nor moral integrity” (Schoenfeld, 2006a, p. 20). He clarified his position in a follow up to the Clearinghouse’s response to his accusations:

Let me be clear about the stakes involved in this case. The issue here is the suppression of a report that challenges the scientific underpinnings of the current federal policy agenda. The WWC website claims that WWC was funded by the government “so that you know what the best scientific evidence has to say.” The evidence suggests otherwise (Schoenfeld, 2006b, p. 23).

Concerns about “problems ranging from severely flawed data to inappropriate methods, to broad conclusions not supported by the evidence provided” (Welner & Molnar, 2007, 6th paragraph) led the Education Policy Research Unit at Arizona State University and the Education and the Public Interest Center at the University of Colorado to create the Think Tank Review Project. Of the 13 reports reviewed by the Center in 2006, “only one, maybe two, could be considered to have even minimally passed expert muster” and “the ideological beliefs of the authors (and their think tanks) appear to have distorted the methods used, shaped the literature reviewed, and determined the results and recommendations of the reports” (Welner & Molnar, 2007, 10th paragraph).

In 2007, the Think Tank Review Project created the Bunkum Awards in Education which identified the worst reports reviewed in 2006. The “Damned Lies

Award for Statistical Subterfuge” was shared by the Manhattan Institute for two evaluations of Florida’s Program to End Social Promotion and by the Program on Education Policy and Governance at Harvard University for a study that attempted to refute the findings of two earlier studies on public versus private school student achievement (Welner & Molnar, 2007, 14th paragraph). [It should be noted that the authors who reviewed the Harvard University study for the Think Tank Review Project (Lubienski & Lubienski, 2006) were also the authors of one of the earlier studies Harvard attempted to refute.]

Summary

Youth service providers who wish to evaluate their programs operate in a climate of an increased emphasis on performance and outcomes, decreased availability of funds, and poor track records of positive outcomes. Many do not have adequate resources or the capacity to conduct effective evaluations. Evaluators who work with youth programs may differ in ideology and may disagree on appropriate evaluation methodologies. These strains will influence procedural and interpersonal aspects of the working relationship between evaluators and their youth service provider clients.

The Practice of Evaluating Youth Programs and Services

Despite the large numbers of problems that have been identified with evaluation results and the calls for improving evaluation processes, the literature is surprisingly limited on the practice of evaluation and insights for providing solutions. Henry and Mark (2003, quoted in Carman, 2007, p. 61) noted that “We know remarkably little about how evaluation is being practiced, why it is being practiced, by whom, and to what effect....Because there is a serious shortage of rigorous systematic evidence.”

Jarosevich, Essenmacher, Lynch, Williams, & Doino-Ingersoll (2006, pp. 12-13) surveyed members of the American Evaluation Association's Independent Consulting Topical Interest Group (TIG) and identified the services most commonly provided to clients. The top seven services were survey/questionnaire development, report writing, qualitative analysis, development of data collection tools, focus groups, evaluation frameworks/logic models, and statistical analysis. They also identified 13 industries in which TIG members provided evaluation services including seven that serve transition-age youth: nonprofit organizations, community-based organizations, social service agencies, K-12 education, health care, postsecondary education, and technical/vocational education (p. 13). Although the quantitative/qualitative e-mail survey of the AEA Independent Consulting TIG may not have had a representative sample, it still provided useful exploratory data on the practice of evaluation by professional evaluators.

Taut and Alkin (2003, p. 213) found that "there is not a noteworthy body of research on the impediments to effectively implementing an evaluation," so they made an exploratory investigation of staff perceptions of barriers to evaluation implementation such as trust/fear issues, poor staff/evaluator relationships, poor communication, lack of staff knowledge of evaluation, and previous bad staff experiences with evaluation. Donaldson, Gooler, & Scriven (2002, p. 262) found "a paucity of literature on the antecedents and consequences of evaluation anxiety within the context of program evaluation," so they investigated "XEA" (excessive evaluation anxiety) on the part of evaluation stakeholders in order to help evaluators think about the psychology of evaluation and avoid problems caused by XEA in their evaluations.

This section of the literature review will examine five aspects of the current practice of evaluating youth programs and services: (a) evaluation resources, (b) an emerging evaluation approach known as Youth Participatory Evaluation (YPE), (c) the evaluation plan, (d) the evaluator-client relationship, and (e) factors influencing the evaluator-client relationship.

Evaluation Resources

There is some evidence that the emphasis on evaluation, participatory or otherwise, by funders and others may not be reaching program operators as a clear, coherent mandate or process. Tilson (1986) investigated the practice of evaluation within school-to-work model demonstration projects funded by the federal Office of Special Education Programs. He examined evaluation plan designs as well as attitudes and perceptions of both federal project officers and local project directors toward evaluation practices. Through document reviews and structured, open-ended interviews, he found that both federal project officers and local program directors had limited training or experience with evaluation. Evaluation expertise varied among local projects, with some hiring external evaluators, accessing in-house evaluation resources, or working with an internal evaluator, while others had no evaluation resources other than those they developed themselves.

In a study of 178 non-profit, community-based organizations (in social services, developmental disabilities, and community development fields) in New York state, Carman (2007) found some surprising and conflicting information about how the organizations viewed evaluation and conducted evaluation activities. For example, although a wide variety of reporting, monitoring, regulatory, and management activities

were reported by respondents, only 65% of the organizations surveyed reported that they conducted formal evaluations of their programs. Only one program reported using an external evaluator, and only ten programs had internal evaluation staff. The rest assigned evaluation responsibilities to executive or management staff, board members or committees, or other staff or volunteers.

Carman concluded that (a) funders should shift their focus from accountability reporting to using evaluation and performance data to improve service delivery and (b) “external evaluators are uniquely positioned to help community-based organizations (CBOs) improve evaluation practice” (2007, p. 72). Carman also found that many CBOs “still do not understand evaluation, whereas others simply lack the capacity to design and implement evaluation in helpful and meaningful ways” (p. 72) and that CBOs need to start investing in their own evaluation capacity” because

the benefits of having the technical infrastructure and internal capacity to conduct ongoing, high-quality data collection and analysis that can be used to inform organizational decision making and improve service delivery would be worth the relatively small investment (p. 73).

Although there are gaps in the transfer of information from evaluation research to practice, especially for people without prior evaluation experience, (Davis, Uhl, Barrington, Rowel, Squiers, Sharp, & O’Brien, 2004), a number of evaluation resources are available to youth organizations—and evaluators. Table 2 contains a sampling of resources and publications for program evaluations which provide information on planning and implementing evaluations in order to examine the evidence that the program is producing the results and/or for accountability purposes. Primary topics of the resources are also listed in the table.

Table 2

A Sampling of Program Evaluation Resources

Resource	Evaluation Concepts and Components
Administration for Children & Families. (2003, December). <i>The program manager's guide to evaluation</i> . Washington, DC: U.S. Department of Health and Human Services.	Selecting an evaluator Evaluation plan Making sense of evaluation information Reporting
Bell, S.H. (2001). <i>Improving the evaluation of DOL/ETA pilot and demonstration projects: A guide for practitioners</i> . (Research and Evaluation Report Series 01-A). Washington, DC: U.S. Department of Labor, Employment and Training Administration.	Choosing the right evaluation type Formative evaluation Descriptive evaluation Operational evaluation Outcome evaluation
Bond, S.L., Boyd, S.E., Rapp, K.A., Raphael, J., & Sizemore, B.A. (1997). <i>Taking stock: A practical guide to evaluating your own programs</i> . Chapel Hill, NC: Horizon Research, Inc.	Self evaluation The evaluation process Summative evaluation Measuring progress Evaluation plan Evaluation report
Chinman, M., Imm, P., & Wandersman, A. (2004). <i>Getting to outcomes 2004: Promoting accountability through methods and tools for planning, implementation, and evaluation</i> . (Technical Report TR-101-CDC). Santa Monica, CA: RAND Corporation.	Accountability Logic model Process evaluation Outcome evaluation Continuous improvement strategies Sustainability
Community Research Associates, Development Associates, & Development Services Group, Inc. (2001) <i>Disproportionate minority confinement technical assistance manual</i> . (2 nd ed., 2 nd printing). Champaign, IL, Arlington, VA, & Bethesda, MD: Authors	Evaluation planning Process evaluation Outcome evaluation Experimental/Quasi/Non-experimental designs Evaluating systems change

Resource	Evaluation Concepts and Components
<p>Friedman, M. (2005). <i>Trying hard is not good enough: How to produce measurable improvements for customers and communities</i>. Victoria, BC: Trafford Publishing.</p>	<p>Results Accountability Thinking process & action plan Population Accountability Performance Accountability</p>
<p>Hatry, H., van Houten, T., Plantz, M.C., & Taylor, M. (1996). <i>Measuring program outcomes: A practical approach</i>. (12th printing). Alexandria, VA: United Way of America.</p>	<p>Outcome measurement Logic model Internal and external uses of outcome findings</p>
<p>Innovation Center for Community and Youth Development & the Institute for Applied Research in Youth Development at Tufts University. (n.d.) <i>Planning and Evaluation Research Center</i> located at http://www.evaluationtools.org</p>	<p>Self evaluation Program planning and evaluation cycle Logic model Evaluation plan Community engagement Youth involvement in evaluation</p>
<p>Innovation Network, Inc. (2005a). <i>Evaluation plan workbook</i>. Washington, DC: Author.</p>	<p>Evaluation plan Logic model Implementation evaluation Outcomes evaluation</p>
<p>McNamara, C. (2007b). <i>Basic guide to program evaluation</i>. Minneapolis, MN & Toronto, Canada: Authenticity Consulting, LLC.</p>	<p>Program evaluation plan Goals-based evaluation Process-based evaluation Outcomes-based evaluation Pitfalls to avoid</p>
<p>McNamara, C. (2007a). <i>Basic guide to outcomes-based evaluation for nonprofit organizations with very limited resources</i>. (Free Management Library) Minneapolis, MN & Toronto, Canada: Authenticity Consulting, LLC.</p>	<p>Outcomes evaluation Evaluation planning</p>

Resource	Evaluation Concepts and Components
National Collaborative on Workforce and Disability for Youth. (2003). <i>Evaluation guide for ODEP youth with disabilities grantees</i> . Washington, DC: Institute for Educational Leadership.	Self assessment based on Baldrige Criteria for Performance Excellence Participatory evaluation strategy
National Training Support Center. (n.d.) <i>ACRN evaluation workbook</i> . Washington, DC: Author.	Nine step evaluation process: Define purpose & scope, Specify evaluation questions, Specify evaluation design, Create data collection action plan, Collect data, Analyze data, Document findings, Disseminate information, Feedback to Program Improvement <i>JCSEE Program Evaluation Standards</i>
Shumer, R. (2007). <i>Youth-led evaluation: A guidebook</i> . Clemson, SC: National Dropout Prevention Center/Network, College of Health, Education, and Human Development, Clemson University.	Doing an evaluation Developing a plan Observations, interviews, surveys, focus groups, participant observations Resources
W.K Kellogg Foundation. (1998, January). <i>W.K. Kellogg Foundation evaluation handbook</i> . Battle Creek, MI: Author.	Context evaluation Implementation evaluation Outcome evaluation Logic Model Stakeholder evaluation team Communicating and utilizing findings/results

Most of the resources in Table 2 include evaluation planning or some kind of explicit evaluation process and describe different types of evaluation or the specific type of evaluation proposed. Not all use the same language to describe evaluation. Some have

unique features such as the *Planning and Evaluation Research Center's* youth involvement model; Friedman's "Results Accountability" model with its emphasis on accessible, shared language; and the National Collaborative on Workforce and Disability for Youth's (NCWD/Y) evaluation assessment based on the Baldrige Criteria for Performance Excellence.

There is still fear and confusion among program operators and staff when it comes to evaluation (Taut & Alkin, 2003; Donaldson, Gooler, & Scriven, 2002) despite the fact that there are numerous evaluation resources at hand. Many program operators conduct evaluations as part of grant requirements (Weiss, 1998a, pp. 22-23). Confusion may result because of unfamiliar evaluation terminology and the plethora of evaluation types and purposes. Problems with making sense of the options and selecting the most appropriate evaluation approach may explain why some resources discuss who should conduct the evaluation, such as evaluation teams and/or an external evaluator (Administration for Children and Families, U.S. Department of Health and Human Service 2003; Weiss, 1998a; McNamara, 2007a and b).

Youth Participatory Evaluation

Youth participatory evaluation (YPE) is a new and emerging approach that is slowly gaining currency in the evaluation field. Kim Sabo Flores, one of the leaders of new field, explains

The practice of engaging young people in evaluation has emerged out of a confluence of diverse constructs and practices, most notably, community development and action research, participatory evaluation, and positive youth development. It has been further fueled by the signing and ratification of the 1989 United Nations Convention on the Rights of the Child (CRC). (Sabo, 2003, p. 1)

Flores explains that youth have a right to be involved in the decisions that affect them and that YPE is important for youth development, the field of evaluation, programs that serve youth, and the community in which the youth live (Sabo, 2003). YPE is a logical extension of youth initiatives such as the Youth Leadership Network and the National Youth Leadership Forum which provide opportunities for youth to gain leadership and career skills, as well as the disability community's "Nothing about us without us" (Charlton, 1998) advocacy and independent living movement for adults and youth with disabilities.

Involving youth as active participants in an activity such as evaluation may concern some community members and evaluators and may be the reason that YPE is not yet a common practice. One explanation is that "concrete models of practice" are needed before evaluators and community members will embrace YPE (London, Zimmerman, & Erbstein, 2003). Another is that concern that the benefits of YPE may be outweighed by the challenges, such as youth frustration when they are not treated seriously or their evaluation results are not used (Checkoway & Richards-Schuster, 2003).

Fortunately, progress is being made as evidenced by resources such as *Youth-Led Evaluation: A Guidebook* (Shumer, 2007), a guide and workbook developed for youth that introduces them to evaluation, the evaluation plan, common data collection methods such as interviews, surveys, and focus, and related resources. It is anticipated that as the research base builds and more training becomes available, concerns about adopting YPE will abate (Checkoway & Richards-Schuster, 2003).

The Evaluation Plan

In discussing evaluation design, Patton (1990, p. 162) says “...*there are no perfect research designs*. There are always trade-offs. These trade-offs are necessitated by limited resources, limited time, and limits on the human ability to grasp the complex nature of social reality.” Difficulties with the design and implementation of program evaluation plans have been documented in a number of contexts (Howell & Yemane, 2006; Jaycox, McCaffrey, Ocampo, Shelley, Blake, Peterson, Richmond & Kub, 2006; Renger, 2006). These difficulties may surface in the original grant proposal, during program implementation, or in the final reporting process. As a result, the effectiveness of the evaluation may be impaired and the desired results may not be attained and/or documented.

The literature suggests that a well-designed and implemented evaluation plan (Howell & Yemane, 2006), based on “clear and well-articulated questions” (Weiss, 1998a, p. 81), is key to an effective evaluation. Weiss (1998a, p. 87) further states that “As questions refer to the content of the study, design is its structure....Design follows questions.”

The funder may describe the type of program evaluation desired in detailed resources such as those cited previously or may provide general guidelines in grant documents. For example, a Solicitation for Grant Applications (SGA) of Innovative Demonstration Grants for Youth With Disabilities from the Office of Disability Employment Policy, U.S. Department of Labor, stated that

proposals must demonstrate how the grantee would develop, implement, evaluate, and disseminate new or improved approaches to the youth programs” ... , [u]se rigorous quantitative or qualitative evaluation methods and data; [e]valuate the model by using multiple measures of results to determine the effectiveness of the

model and its components or strategies for continuous program improvements; [and p]roduce detailed procedures and materials that would enable others to replicate the model (Kuss, 2002, p. 43171).

ODEP further stated that factors to be considered in evaluating the grantee's project design and evaluation plan included

“the extent to which performance feedback and continuous improvement are integral to the design of the proposed project,...the extent to which the methods of evaluation... are thorough, feasible, and appropriate...; provide for examining the effectiveness of project implementation strategies;...include the use of objective performance measures that are clearly related to the intended outcome of the project and will produce quantitative and qualitative data; ...provide information to other youth programs about effective strategies; and measure...program results and satisfaction of youth with disabilities” (Kuss, 2002, p. 43173).

Using the ODEP Solicitation for Grant Applications as an example, the grantee must develop an evaluation plan that will include—at a minimum—quantitative and qualitative methods and data; multiple objective performance measures for program effectiveness, continuous improvement, and youth satisfaction; and a dissemination plan for sharing information with other youth programs. These requirements describe formative, summative, outcome, and process evaluations. But more decisions need to be made: What program model will be used? Who will conduct the evaluation? What will the evaluation methodology be?

As noted previously, each evaluation has unique purposes, methodologies, resources, and intended uses. A simple framework for developing an evaluation plan is the RUFDATA model developed in England as “a practical approach to evaluation planning” (Saunders, 2000, p.7). Using the RUFDATA framework (Saunders, 2000, p. 18) and the ODEP SGA guidelines as an illustration, potential questions and answers for a program evaluation are presented in the next table.

Table 3

Possible Evaluation Planning Questions using the RUFDATA Model and ODEP SGA Guidelines

What are our Reasons and Purposes for evaluation?

- To meet funder program evaluation requirements
- To evaluate the youth program model and strategies including rigorous and appropriate evaluation methods and data as well as detailed procedures and materials (Kuss, 2002).

What will be the Uses of our evaluation?

- To share information on effective strategies and models with other youth programs who may wish to replicate them
- To inform stakeholders of program progress and outcomes
- To inform the funder of program progress and outcomes
- To share successes with the media for public relations purposes

What will be the Foci of our evaluation?

- Formative evaluation for continuous program improvement (Hanson, 1998)
- Process evaluation to provide program replication information (Hanson, 1998)
- Outcome evaluation to determine effectiveness of the program model and strategies (Hanson, 1998)
- Summary evaluation to assess performance measures for accountability and decision-making (Hanson, 1998)

What will be the Data and Evidence for our evaluation?

- Participant satisfaction surveys
- Participant enrollment, attendance, and participation rates; assessment scores; numbers of work experience sites, etc.
- Numbers of participants who graduate from high school, enroll in post-secondary education, and become employed

Who will be the Audience for our evaluation?

- Other youth program operators
 - Stakeholders such as program staff, board members, partners, community members...
 - Program participants
 - The media/The general public
-

What will be the Timing for our evaluation?

- Performance monitoring and data will be reported quarterly to the funder
- Continuous improvement data will be reported as needed, but at least quarterly, to the staff and program manager
- Program data will be reported annually to the board
- Final outcomes and results will be reported at the end of the grant cycle to funder and stakeholders

Which Agency should be conducting the evaluation?

- A participatory evaluation model will be used where an internal or external evaluator guides the technical research and convenes stakeholders regularly to discuss progress and implications (Weiss, 1998a, p. 99-100)

(Kuss, 2002; Saunders, 2000).

Note: The RUFDATA framework is used solely as an example; whether it is appropriate for a given evaluation or program must be determined by program stakeholders and evaluators.

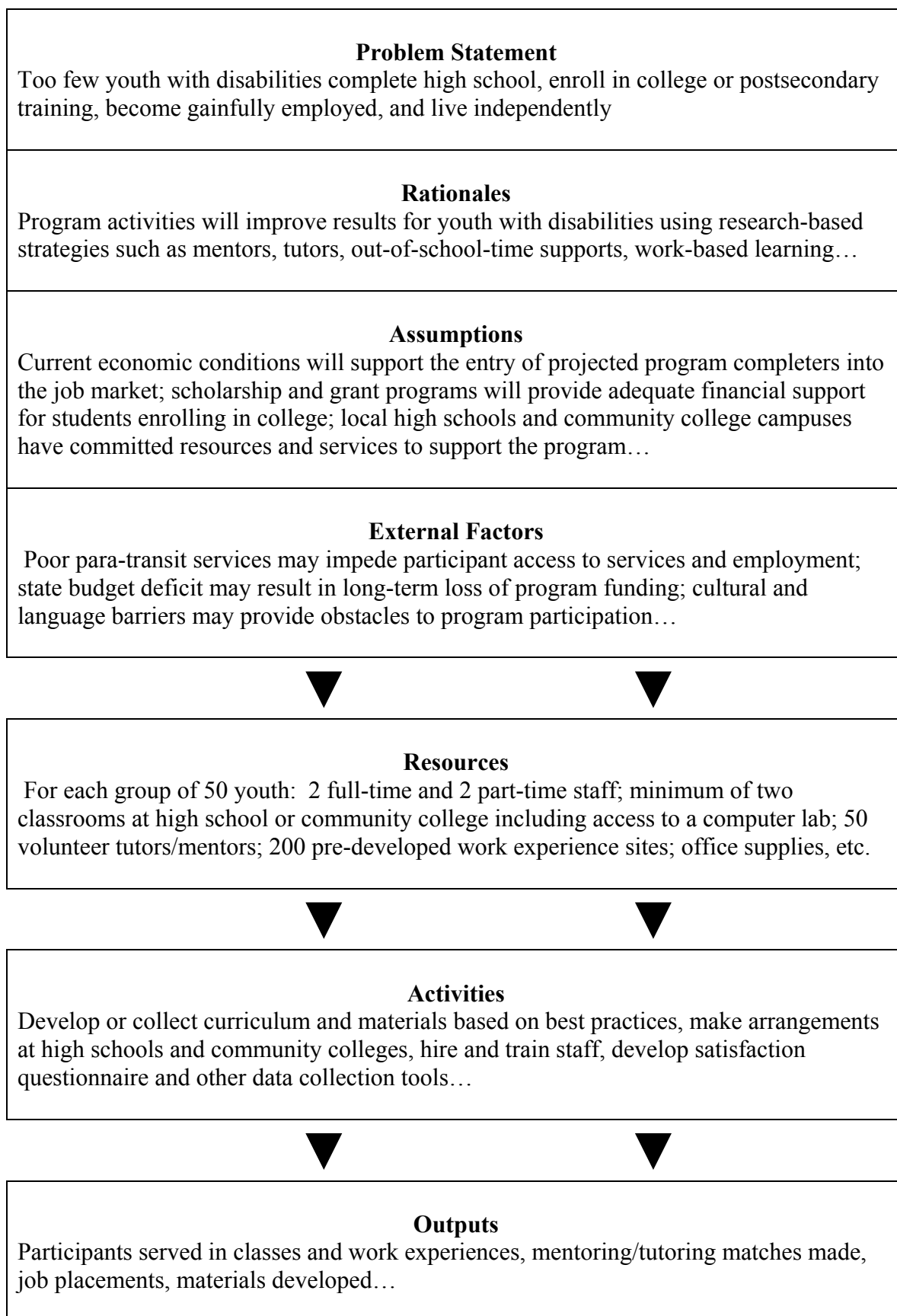
After the evaluation planning questions are developed and answered, a logic model and/or theory of action or change—required by many funders—can be developed to explain or diagram how the program or intervention will work. Commonly used models are the United Way Program Outcome Model (Hatry, van Houten, Plantz, & Taylor, 1996), the W.K. Kellogg Logic Model (2001), and the Innovation Network Logic Model (2005b). However, practitioners may find these models to be confusing and difficult to implement (Cecil, 2004) as many have had limited or no formal training and/or knowledge in designing and implementing program evaluations (Tilson, 1986). They also may have different attitudes and understandings of the purposes of evaluation (Carman, 2007; Tilson, 1986).

Cecil (2004) identified a number of facilitating factors and barrier factors in transferring the United Way Program Outcome Model to human service provider

programs. Barriers included model deficiencies, lack of resources, lack of evaluation capacity, inadequate guidance, and leadership issues (Cecil, 2004, pp. 115-133).

Facilitating factors included proactive leadership, adequate guidance, organizational resources and capacity, and a user-friendly model (Cecil, 2005, pp. 83-100).

Commonly occurring logic model components are illustrated in Figure 1 using the Innovation Network model and a hypothetical response to the ODEP grant requirements.



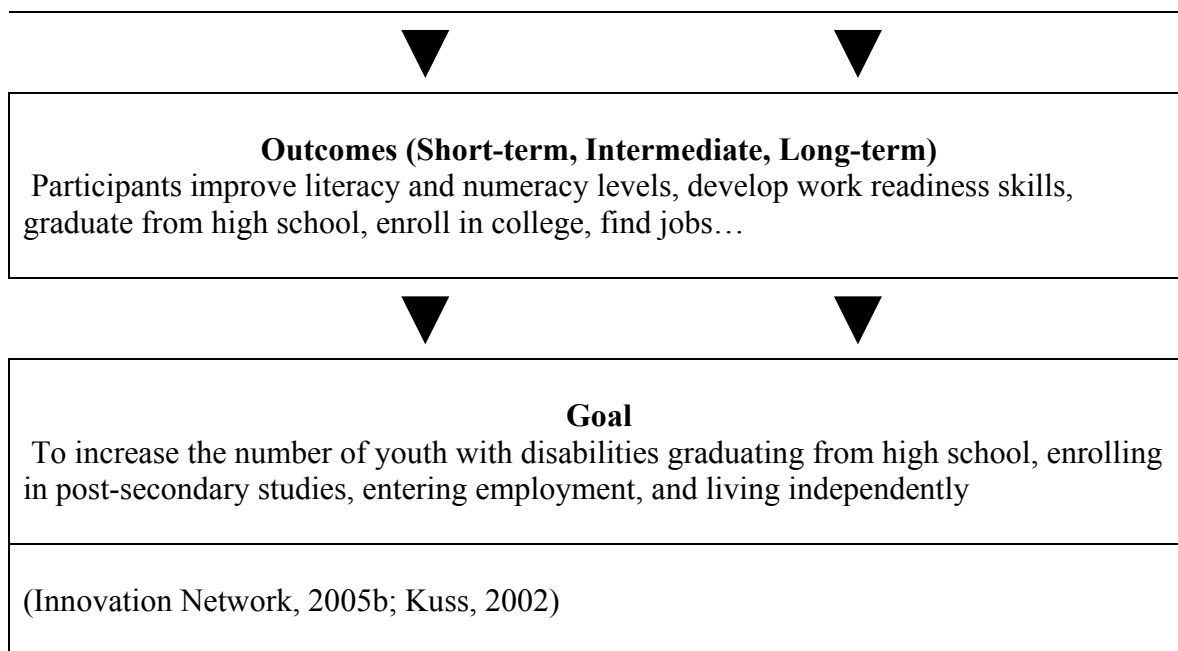


Figure 1. Innovation Network Logic Model for Possible ODEP Grantee

Once the logic model and/or theory of action are developed, attention can be turned to the methodology and other components of the evaluation plan. As noted previously, the ODEP SGA required formative, summative, outcome, and possibly process (or operational) evaluations. Quantitative (e.g., numbers of participants served, completed, graduated from high school, placed in employment) and qualitative (subjective information, e.g., participants likes and dislikes about the program, changes participants would make in the program) data collection for outcome and continuous improvement evaluation were also specified. The challenge is to design an evaluation plan that will “maximize what is learned...while holding costs and intrusiveness of the research to a minimum. Different approaches are recommended in different situations, based on state-of-the-art research techniques and recognition of the practical constraints facing ‘real world’ field studies” (Bell, 2001, p. 1).

If the funder does not specify an evaluation methodology, the program operator will determine the methodology based on program goals (“Reasons and Purposes” in the RUFDATA Model). Bell (2001, pp. 13-17) developed a chart which maps program goals and evaluation methodology and noted that the program operator may need to drop some of the program goals in order to develop a manageable evaluation methodology. Program characteristics, available resources, and other factors may also cause changes in program goals.

Some researchers (Weiss, 1998a; Bell, 2001) recommend an evaluability assessment to determine if the evaluation study “can be understood conceptually, implemented faithfully, measured reliably, and interpreted unambiguously” (Bell, 2001, p. 31). The evaluability assessment should answer the following questions:

1. Has the proposed intervention been specified in enough detail, conceptually, to allow for unambiguous statements about its nature and relationships to other distinctive policies, once research results are available?
2. Does the intervention as planned have a reasonable chance of being “implementable”—that is, is it capable of execution in a real-world setting in rough conformity to its intent?
3. Will enough observational information about the intervention be available to allow for descriptions of its nature and judgments as to its success or failure, strengths, and weaknesses? (Bell, 2001, p. 31).

“If the answers to any of the above questions is ‘no,’ the next step is to identify changes that would right the situation, or to determine that the evaluation should not go forward” (Bell, 2001, p.32). Program operators will also need to decide whether and when to bring in an external evaluator to manage the evaluation, provide technical assistance, or manage the evaluation in-house.

Once the evaluation plan has been developed, program operators will need to address specific challenges in order to implement the plan—part of the evaluation process

that many resources do not address. A sampling of implementation questions that will need to be answered may include:

1. When and how will existing and new staff be trained or updated on new program and evaluation procedures and reporting requirements, particularly if rigorous evaluation methodologies such as experimental or quasi-experimental designs are specified?
2. Are existing data collection systems capable of collecting and compiling the necessary program data including baseline, progress, and outcome data—or will a new data system need to be developed?
3. What new forms and reports will need to be developed?
4. Is there an existing continuous improvement system in place or will a new one be needed?
5. Does the evaluation methodology address legal issues such as client confidentiality and informed consent?
6. What about identifying and addressing ethical issues and practices?
7. Are evaluation responsibilities clearly delineated among staff such as who will be responsible in each site for managing the process and who will have oversight responsibilities for compiling and sending reports to the funder and stakeholders?
8. How will promising practices, effective strategies, lessons learned, resources developed, and/or economic analyses (cost benefit, cost effectiveness, return on investment, etc.) be identified, calculated, and packaged?

9. Is a dissemination structure in place for sharing evaluation progress and results, or will a new one need to be developed?
10. Can management and program staff quickly respond to external issues impacting the evaluation plan such as reductions in funding, changes in funder ideology or priorities, or disruptions in the local economy?
11. Can management and program staff quickly respond to internal issues impacting the evaluation plan such as staff turnover, scheduling conflicts with training facilities, computer or data analysis problems, and so forth.

Assisting the client in identifying and answering implementation questions is often a critical part of the evaluator's role.

The Evaluator-Client Relationship

To summarize, the context in which the evaluator-client relationship takes place is one where

- funders are calling for improved program performance but the amount of funding is becoming increasingly tight,
- youth programs have difficulty showing that their program interventions are effective, although it is not clear whether the program or the evaluation (or both) is the issue,
- there are numerous challenges to implementing an effective youth program evaluation such as ideological biases, opposing views on methodologies, factional disagreements, and technical limitations on the part of youth program providers,
- knowledge of how evaluation works in practice is lacking

- evaluation resources are numerous, but less so on the topic of youth program evaluation,
- youth participatory evaluation is beginning to emerge,
- designing and implementing an effective evaluation plan can be problematic and anxiety-producing for the client, and
- the literature on the evaluator-client relationship focuses more on the procedural aspects of the relationship rather than the interpersonal aspects, as described in the following section.

Members of the American Evaluation Association who provide program evaluation services are guided by two primary documents that describe professional practice and, in so doing, describe professional interactions between evaluators and their clients or stakeholders. They are the *American Evaluation Association Guiding Principles* (AEA, 2004) and *The Program Evaluation Standards* of The Joint Committee on Standards for Educational Evaluation (1994).

The *American Evaluation Association Guiding Principles* were developed to “proactively guide the behaviors of professionals in everyday practice” and to “inform evaluation clients and the general public about the principles they can expect to be upheld by professional evaluators” (AEA, 2004, Preface, Section C). The five principles are

1. Systematic Inquiry: Evaluators conduct systematic, data-based inquiries.
2. Competence: Evaluators provide competent performance to stakeholders.
3. Integrity/Honesty: Evaluators display honesty and integrity in their own behavior, and attempt to ensure the honesty and integrity of the entire evaluation process.
4. Respect for People: Evaluators respect the security, dignity and self-worth of respondents, program participants, clients, and other evaluation stakeholders.
5. Responsibilities for General and Public Welfare: Evaluators articulate and take into account the diversity of general and public interests and values that

may be related to the evaluation. (AEA, 2004, The Principles, Items A, B, C, D, & E.)

The Guiding Principles describe a number of interactions between evaluator and client such as discussions of evaluation approaches, limitations, and methods; honest negotiations of evaluation plans and procedures; full disclosure of potential conflicts of interest or values; respecting differences such as culture or disability among stakeholders; inclusion of all stakeholder perspectives in evaluation reports; and discussion and resolution of conflicts between stakeholders or stakeholder needs (AEA, 2004). The Principles also address professional ethics, informed consent, confidentiality, and other procedural issues.

The Program Evaluation Standards of The Joint Committee on Standards for Educational Evaluation complement and reinforce *The Guiding Principles* and describe standards in four areas:

1. Utility standards to ensure that an evaluation will serve the information needs of intended users.
2. Feasibility standards to ensure that an evaluation will be realistic, prudent, diplomatic, and frugal.
3. Propriety standards to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results.
4. Accuracy standards to ensure that an evaluation will reveal and convey technically adequate information about the features that determine worth or merit of the program being evaluated (The Joint Committee on Standards for Educational Evaluation, 1994, Summary of the Standards.)

Client-evaluator interactions described in *The Program Evaluation Standards* include working with interest groups to ensure their support, respect for stakeholders and human subjects, and responsiveness to stakeholder concerns and issues. *The Standards* also address procedural issues such as evaluation reports, written agreements, evaluation

findings, financial responsibility, evaluation documentation and procedures, reporting, and evaluation of the evaluation.

The *Encyclopedia of Evaluation* states

One difference between evaluators and researchers is that evaluators are hired by clients. Many clients are new to evaluation or have inappropriate views of the evaluator's role and the nature of evaluation work. For this reason and many others, the evaluator-client relationship is fraught with potential ethical conflicts (Mathison, 2005).

The *Encyclopedia* mentions the AEA *Guiding Principles* and *The Program Evaluation Standards* and describes several aspects of evaluation with ethical implications such as the importance of a written agreement that describes evaluator and client roles during the evaluation, legal and ethical concerns related to data confidentiality and management, the reporting of results, and differences in the way evaluators view ethical issues. For example, she notes that "...internal evaluators are less likely than external evaluators to see problems through an ethical lens" (Mathison, 2005, p. 131).

Malloy and Yee (2006, p. 67) found that

...our success in the field of evaluation depends just as much on cultivating effective client relationships as it does on our skills and expertise. We have also learned that when we approach our client relationships as collaborative partnerships, we produce more effective evaluations and enjoy more professionally rewarding experiences.

Malloy and Yee (2006) also identified a number of literature- and experience-based conditions that must be met in order for the client-evaluator partnership to be effective. The conditions include "complementary needs and assets" (p. 69), "compatible goals" (p. 70), "trust" (p. 70), "explicit decision-making process and governance structures" (p. 72), "an accountability plan" (p. 73), "effective leadership" (p. 73), and regular review of the positive and negative impacts of the partnership (p. 74).

Wandersman & Snell-Johns (2005) also mention the *Guiding Principles* and “*Joint Standards*” in discussing “the appropriate nature of the evaluator-client relationship” (p. 424) which varies based on the circumstances of each individual evaluation. A particular concern was whether “intentionally amiable relationships between evaluators and program stakeholders” in empowerment evaluations create or reduce bias in the evaluation (p. 424-425). Empowerment evaluations build the evaluation capacity of program stakeholders so that evaluation can become part of program operation; consequently the evaluator’s role changes over time as stakeholders take over progressively more evaluation responsibilities.

Weiss (1998a, p. 112) notes that “Whereas the traditional role for the evaluator has been that of dispassionate observer and reporter on things as they are, current thinking has introduced a range of possible alternatives.” Rossi, Lipsey & Freeman (2004, p. 402) add that

Whether evaluators are insiders or outsiders, they need to cultivate clear understandings of their roles with sponsors and program staff. Evaluators’ full comprehension of their roles and responsibilities is one major element in the successful conduct of an evaluation effort.”

Patton (2008, pp. 210-211) is more explicit and provides a “menu” for “matching primary users [of evaluation results], primary evaluator roles, dominant style of evaluator, most likely evaluation purpose, and primary evaluator characteristics affecting use [of evaluation results]” when negotiating evaluation terms with potential clients.

A number of authors have described evaluator roles. Some evaluator roles are formal such as judge of program merit or worth (Scriven, 1991; Stufflebeam, 1994; Patton, 1997), auditor, inspector, or compliance/accountability evaluator (Barkdoll, 1980; Patton, 1997) or researcher/scientist (Barkdoll, 1980; Weiss, 1998a; Patton, 1997). Some

roles are collaborative such as evaluation facilitator or guide (Weiss, 1998a; Patton, 1997; Wiltz, 2005); team member with evaluation perspective or internal advisor; (Patton, 1997, Wiltz, 2005); collaborator or partner (Barkdoll, 1980; Weiss, 1998a; Patton, 1997; Wiltz, 2005); or critical friend (Rallis & Rossman, 2000; Weiss, 1998a). Other roles take an organizational development approach such as program improvement consultant (Patton, 1997); educator or capacity builder (Morabito, 2002. Wiltz, 2005; Weiss, 1998a); organizational development consultant or change agent (Owen, Lambert & Stringer, 1994; McClintock, 2003; Wiltz, 2005); or empowerment facilitator who trains program administrators or staff to conduct evaluations on their own (Weiss, 1998a; Patton, 1997). Still other evaluator roles may include supporter of a cause or advocate for social change (Patton, 1997; Morabito, 2002) or program advocate (Wiltz, 2005).

Factors Influencing the Evaluator-Client Relationship

The client-evaluator relationship is a construct with both procedural and interpersonal components, both of which are critical for planning and implementing effective evaluations—and both of which are influenced by internal and external forces. External forces influencing the client-evaluator relationship may include funder requirements such as specific evaluation methodologies or plan components, changes in governing laws or regulations, or unanticipated challenges in planning or implementing the evaluation. Internal forces may include stakeholder resistance to evaluation planning and/or implementation, lack of program staff familiarity with evaluation methodologies and resource materials, problems with data collection and management, and stakeholder concerns about how the evaluation results will be used.

The literature identifies a number of factors that influence the evaluator-client relationship, although most are not discussed in the context of that relationship. For example, it is likely that the interpersonal and procedural dynamics will be very different for evaluator roles such as judge, compliance monitor, collaborator, teacher, organizational consultant, or advocate. The reason for hiring an evaluator or evaluation consultant will set the stage for the evaluator-client relationship. Hiring an evaluator may be a funder requirement (Kegeles, Rebchook, & Tebbetts, 2005; Rossi, Lipsey & Freeman, 2004; Manela & Moxley, 1999). It may also be due to a lack of organizational capacity, external requirements such as accreditation, or a decision by the youth organization's board of directors or senior management (Thayer, 2006).

The focus of the evaluation may also have an impact on the evaluator-client relationship. Hanson (1998, p. 44) identified five evaluation foci that have different purposes and potential repercussions for the client:

1. Needs assessment to verify and map the extent of a problem
2. Formative evaluation to improve and refine the program
3. Process evaluation to aid program implementation and replication
4. Outcome or impact evaluation to assess program effectiveness in producing change
5. Summative evaluation to assess program quality/impact for accountability and decision-making purposes

As discussed previously, a well-designed and implemented evaluation plan (Howell & Yemane, 2006) is necessary for effective evaluation. How well the plan is designed and implemented will have an effect on the evaluator-client relationship. Jackson (2004) identified eight components that should be discussed and agreed upon during the evaluation planning process: program readiness for evaluation; contract terms; the focus of the evaluation; the evaluation methodology, field tests of instruments,

materials, or processes; the methodology or research design; data collection, editing, and reporting; and reporting of the evaluation results.

Even a well-designed and implemented plan may encounter challenges during the course of the evaluation which may stress the evaluator-client relationship. The literature identifies a number of challenges:

1. Rigid organizational procedures or settings (Taut & Alkin, 2003; Weiss, 1998a; Cecil, 2004).
2. Inadequate or problematic data collection, management, analysis, or reporting (Donaldson, Gooler & Scriven, 2002; Taut & Alkin, 2003; Kegeles, Rebhook, & Tebbetts, 2005; Stoecker, 2007).
3. Disagreement/debate over who should be conducting the evaluation (Stoecker, 2007).
4. Lack of cooperation, collaboration, or communication among staff (Weiss, 1998a; Donaldson, Gooler, & Scriven, 2002).
5. Limited methodological capacity of staff (Hasenfeld & English, 1974; Kegeles, Rebhook, & Tebbetts, 2005; Stoecker, 2007).
6. Lack of staff resources such as time and funding (Gilliam, Davis, Barrington, Lacson, Uhl, & Phoenix, 2002; Hasenfeld & English, 1974; Taut & Alkin, 2003; Kegeles, Rebhook, & Tebbetts, 2005; Stoecker, 2007).
7. Unexpected/unanticipated evaluation results or consequences (Donaldson, Gooler, & Scriven, 2002; Taut & Alkin, 2003).
8. Friction between youth organization and funder (Stoecker, 2007; Gilliam, Davis, Barrington, Lacson, Uhl, & Phoenix, 2002).

9. Misapplication of theories such as constructionism or post-modernism (Weiss, 1998a; Thayer, 2006).

10. Evaluation results not used (Donaldson, Gooler, & Scriven, 2002).

Threats to the validity of the evaluation may also be encountered and may include

1. Outside forces affected intervention or results (Jackson. n.d; Taut & Alkin, 2003; Weiss, 1998a; Mulroy & Lauber, 2004, Peterson & Randall, 2006).

2. Interventions mutated from the original design (Jackson. n.d; Peterson & Randall, 2006).

3. Control groups were not comparable (initially or over time) (Jackson. n.d).

4. Youth reacted to please or displease program operators (halo/cloud effects) (Jackson. n.d).

5. Program interventions migrated to the control group (Jackson. n.d).

6. Youth became sensitized to intervention measures over time (Jackson. n.d).

7. Maturation of participants clouded intervention results (Jackson. n.d).

8. The evaluation process caused disruptions that impacted results (Jackson. n.d; Weiss, 1998a; Peterson & Randall, 2006).

9. Measurements became unstable or changed over time (Jackson. n.d).

Reports of the usefulness of the evaluation and the number of audiences with whom the results are shared may provide additional insights into the evaluator-client relationship. A well-implemented and useful evaluation will result in improvements in stakeholder focus, ownership, and skills (Cecil, 2004; Gilliam, Davis, Barrington, Lacson, Uhl & Phoenix, 2002, Thayer, 2006); organizational capacity, impact, viability, and credibility (Cecil, 2004; Thayer, 2006); program implementation and results (Thayer,

2006; Gilliam, Davis, Barrington, Lacson, Uhl, & Phoenix, 2002); and use of evaluation results (Gilliam, Davis, Barrington, Lacson, Uhl, & Phoenix, 2002). Evaluation results may be shared with program sponsors and/or funders, program directors and staff, program participants, managers of similar programs, policymakers and opinion leaders, social scientists, other evaluators, and the general public (Weiss, 1998a).

Summary

Evidence has been presented that many youth service providers who wish to evaluate their programs are confused about evaluation and lack the knowledge and capacity to conduct effective evaluations. There is not much research on the current practice of evaluation despite numerous guides, texts, web sites, and other resources on the subject. However, it is known that a well-designed and implemented evaluation plan (Howell & Yemane, 2006), based on “clear and well-articulated questions” (Weiss, 1998a, p. 81), is key to an effective evaluation—as is a positive working relationship between the evaluator and client (Malloy & Yee, 2006).

The evaluator-client relationship has procedural and interpersonal components that are influenced by a number of factors. These factors include the resources used to plan and guide the evaluation, the role of the evaluator, the reason the evaluator was hired, the focus or purpose of the evaluation, the evaluation design as described by evaluation plan components, whether youth are involved in evaluation planning and implementation, challenges to evaluation implementation, threats to evaluation validity, and the effectiveness of the evaluation as measured by improvements resulting from the evaluation and audiences with whom the results are shared.

Conceptual Framework for Research Study

The conceptual framework for the study is based on the theoretical frameworks of program evaluation, organizational change, and constructivism and the disparate knowledge, experiences, training, and perspectives of evaluators, youth program staff and stakeholders. Differences in perspectives and understanding form “parallel universes” (a researcher-constructed concept) where one universe is inhabited by program staff/stakeholders and the other by evaluators, as illustrated in Figure 2. External forces and internal organizational and interpersonal dynamics impact the evaluator/client relationship, which should be the bridge between the two universes. The focus of the research study, as reported by professionals in the evaluator universe, is the factors that influence the evaluator-client relationship—positively by merging the parallel universes into a functional whole or negatively by keeping them apart—thereby impacting the effectiveness of the evaluation.

Theoretical Foundations

Program Evaluation Theory: Youth program operators and AEA evaluators have different levels of understanding and knowledge of program evaluation theory as well as varying amounts of experience implementing program evaluations.

Organizational Change Theory: Youth program operators are part of the organization and experience the evaluation change process differently than AEA evaluators who are part of the team leading the evaluation and may or may not be part of the organization.

Constructivist Theory: Youth program operators and AEA evaluators construct individual, unique meanings based on different personal experiences, education, and motivations resulting in different understandings of program evaluation and reactions to implementing the changes required.

Study Goals

- Describe the factors influencing the evaluator-client relationship.
- Provide information on the practice of evaluation to evaluators, funders, and program stakeholders.
- Improve evaluation effectiveness.

	Youth Program Operators		AEA Evaluators	
<p>E X T E R N A L F O R C E S</p>	Evaluation process often initiated for external reasons such as funder or grant requirements	<p>I N T E R N A L F O R C E S</p>	Initiates the evaluation process at the request of the program operator or board.	<p>E X T E R N A L F O R C E S</p>
	Historic distrust of accountability requirements, evaluation goals, and external evaluators		Passion for good evaluation with the goal of improving program and evaluation practices and outcomes	
	Limited knowledge of evaluation terminology, theory, purpose, and process		Extensive knowledge of evaluation terminology, theory, purpose and process	
	Evaluation plan and/or components based on external/funder requirements		Evaluation plan and components based on theory and good practice	
	Focus on implementation of services, not planning/designing/ implementing evaluation		Focus on importance of evaluation design and implementation to improve services and results	
	Share program results and process issues, if at all, with a		Share evaluation results and lessons learned with	

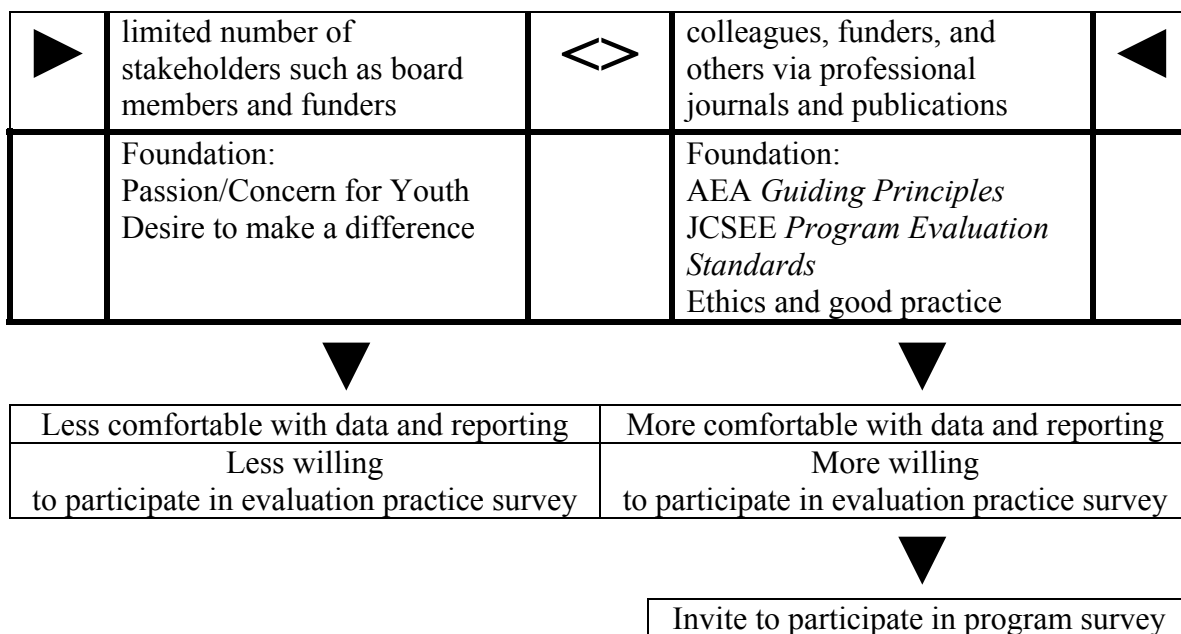


Figure 2. Parallel Universe Conceptual Framework for Research Study

Inferences for the Forthcoming Study

A review of the literature indicates that youth-serving organizations that rely on competitive, discretionary grants to fund their programs or services operate in a landscape of (a) increased scrutiny and demands for accountability and positive results by funders and monitoring agencies (The U.S Office of Management and Budget, n.d.a; Ashby, 2007; Barton & Wilhelm, 2007; Buechel, Keating & Miller, 2007; Executive Office of the President of the United States, Office of Management and Budget, 2007; Nilsen, 2007; Robertson, 2007; Welner & Molnar, 2007; Bellis, 2006; Czerwinski, 2006; Shaul, 2006; Posner & James, 2005; Shaul, 2005; Bellis, 2004; Davis, Uhl, Barrington, Rowel, Squiers, Sharp, & O'Brien, 2004; Shaul, Edwards, & Fucile, 2004; Blank, Melaville, & Shah, 2003; Ekstrand & Rezmovic, 2003; Ekstrand, 2001; Kingsbury, 2000) and (b) on-going problems with data collection and evaluation (Chudowsky, Chudowsky

& Kober, 2007; New Commission on the Skills of the American Workforce, 2007; Robertson, 2007; What Works Clearinghouse, 2007; Bellis, 2006; Branch-Smith, Gray, Fruchter, Hernandez, Joselowsky, Nichols-Solomon, Simmons, & Warren, 2006; Casner-Lotto, 2006; Horn, 2006; National Center for Public Policy and Higher Education, 2006; National Dropout Prevention Center for Students with Disabilities, 2006; Barton, 2005; Curran, 2005; NISS/ESSI, 2005; Shaul, 2005; Carey, 2004; Shaul, Edward, & Fucile, 2004; Bain, 2003; Blank, Melaville, & Shah, 2003; Sum, Harrington, Bartishevich et al, 2003; Gladieux & Swail, 1998).

A multitude of resources is described in the literature for implementing evaluations and improving program outcomes (National Training Support Center, n.d.; Friedman, 2005; The Innovation Network, 2005a & b; Chinman, Imm & Wandersman, 2004; Administration for Children and Families, U.S. Department of Health and Human Services, 2003; National Collaborative on Workforce and Disability for Youth, 2003; Bell, 2001; Community Research Associates, Development Associates, & Development Services Group, Inc., 2001; W.K. Kellogg Foundation, 2001; W.K. Kellogg Foundation, 1998; Bond, Boyd, Rapp, Raphael, & Sizemore, 1997; McNamara, 1997-2007a; McNamara, 1997-2007b; Hatry, van Houton, Plantz, & Taylor, 1996). However, it is not clear how many youth grantees are aware of or use the resources.

The literature also contains numerous descriptions or evaluations of successful or unsuccessful program evaluations and the resulting impacts on programs and results—although youth grantees may not understand or use them (Carman, 2007; Howell & Yemane, 2006; Jaycox, McCaffrey, Ocampo, Shelley, Blake, Peterson, Richmond & Kub, 2006; Renger, 2006; Cecil, 2004; Davis, Uhl, Barrington, Rowel, Squiers, Sharp, &

O'Brien, 2004; Taut & Alkin, 2003; Donaldson, Gooler, & Scriven, 2002). What the literature does not show is how evaluations are done in practice (Henry and Mark, 2003, quoted in Carman, 2007), including examinations of the evaluator/client relationship and internal and external factors that influence the relationship.

It can be inferred from the theoretical framework for the proposed study that professional evaluators and youth program providers have different attitudes, perceptions, knowledge, and understanding regarding evaluation that will impact the development and implementation of evaluation plans for transition programs for youth ages 14-25. It can also be inferred that a number of internal and external factors may create stresses and challenges to effective evaluator/client relationships which may account, in part, for the documented problems with evaluation methodologies and the resulting lack of definitive results for youth transition programs. Finally, it may be inferred that examining and understanding the factors that influence the evaluator/client relationship will consequently improve both the evaluation process and results for transition programs serving youth ages 14-25.

CHAPTER THREE

“Exploratory data analysis can be defined as the examination of data with minimal preconceptions about its structure through which it is hoped that relationships and patterns, at least some of which are unanticipated, will be discovered” (Wainer, 1978, p. 151). In the process, questions or hypotheses will be generated for future analysis (Tukey, 1980; Brilliner, 2002). The purpose of this study was to explore factors influencing the evaluator/client relationship, based on a theoretical model that posits that relationship as a critical element of an effective evaluation, in recent evaluations of programs serving transition age youth (ages 14-25) from the perspective of professional evaluators working with the youth grantees and to generate questions about the findings. Survey data was collected from members of the American Evaluation Association using an electronic, researcher-constructed survey instrument. Designing, validating, and piloting a survey instrument was not a purpose of the study but was necessary because no instrument was found that addressed the study’s research questions. Simple descriptive data analysis techniques were used to analyze the quantitative data, and an inductive coding process was used to analyze the qualitative data. This chapter presents an overview of the research problem, the research purpose and questions, the research design, the population and sampling selection, the survey instrument used, data collection procedures, data analysis methods, and ethical considerations.

Research Purpose and Questions

The literature suggests that a well-designed and implemented evaluation (Howell & Yemane, 2006) and effective evaluator-client relationships (Malloy & Yee, 2006) are necessary for effective evaluation. Therefore, the purpose of the study was to explore

factors affecting the evaluator-client relationship in recent evaluations of programs serving transition age youth (ages 14-25) from the perspective of professional evaluators working with the youth grantees and to generate questions about the findings. The research questions for this study were:

1. What are the characteristics of the organizations whose youth program was evaluated?
2. What are the characteristics of the role of the evaluator or evaluation consultant?
3. What are the characteristics of program evaluation planning and implementation that may impact the evaluator/client relationship?
4. What questions relating to the evaluator-client relationship result from analyzing the data collected in questions 1, 2, and 3?

Overview of Methodology

Members of the American Evaluation Association (AEA) who worked as evaluators or evaluation consultants with organizations serving youth ages 14-25 in the United States were surveyed. The Internet-based survey contained specific questions about the type of youth organization, the role of the evaluator or consultant, challenges to evaluation implementation, threats to evaluation validity, evidence of successful evaluation implementation, and other relevant information. Data analyses included frequency calculations for quantitative responses, reliability tests (Cronbach alpha) for Likert scale questions, and inductive and deductive analysis of open-ended qualitative questions.

Research Design

An exploratory, descriptive research study was implemented employing quantitative and qualitative methods. Participants were solicited from the American Evaluation Association to complete a researcher-constructed survey. The cross-sectional, Internet-based survey was composed of 15 qualitative and three quantitative questions. The literature review found no survey instrument that answered the study's research questions, including that of a recent survey of American Evaluation Association members, so a new survey instrument was designed. Survey questions focused on factors influencing the evaluator-client relationship based on the literature review and the researcher's experiences. (See Appendix B for a question-by-question listing of survey citations.) Figure 3 displays a flow chart describing the steps from survey construction to data analysis employed for this study based on a model developed by Scheuren (2004).

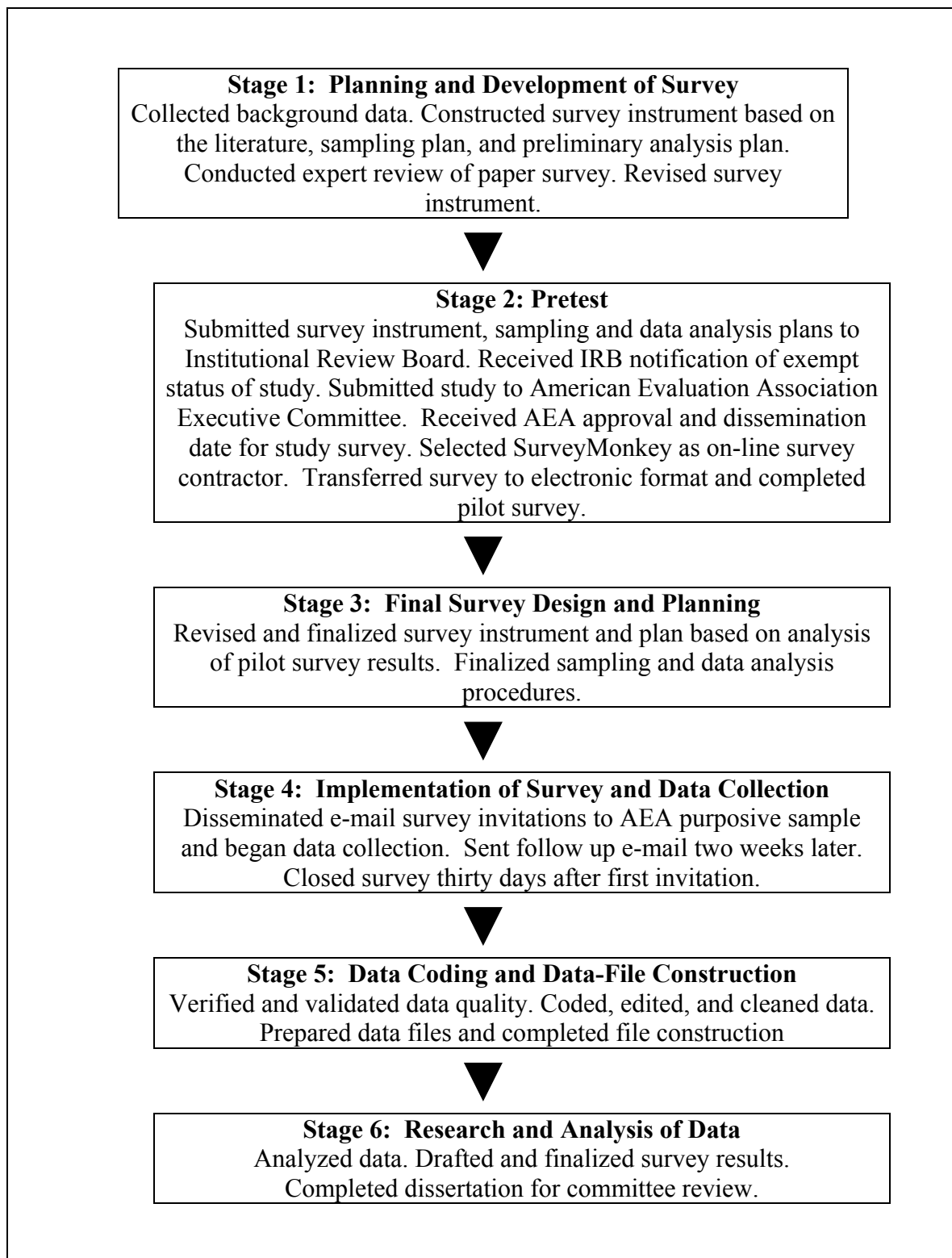


Figure 3. Flow Chart for Methodology

Population and Sample

The study population was composed of members of the American Evaluation Association, (AEA), an association of professional evaluators and related professionals. The association was selected for this study because of its membership's high professional standards, broad range of evaluation experiences, variety of evaluation positions, and demonstrated concerns for promoting valid, responsive evaluation practice. A criterion based, purposeful sample was composed of AEA members who voluntarily responded to the survey invitation, were members of one or more of 12 AEA Topical Interest Groups that were likely to serve youth, and who self-identified as

1. a member of AEA
2. providing evaluation services in the United States
3. acting as an evaluation consultant, provided evaluation technical assistance, or served as an internal or external evaluator
 - a. in the last *five* years
 - b. to organization serving *youth ages 14-25*
 - c. funded by discretionary or competitive *grants*.

However, membership data collected by AEA includes several categories that serve youth organizations but do not identify organizations serving the 14 to 25 age range. This age range also overlaps with some adult programs. Consequently, it could not be determined if the purposive sample was representative of the full population.

Instrumentation

Institutional Review Board and American Evaluation Association research guidelines were followed throughout the research process. The research instrument was

based on the literature describing evaluator roles and other factors influencing the evaluator-client relationship and the experiences of the researcher in providing technical assistance on evaluation to youth service providers. The draft survey instrument was reviewed by five experts in the fields of evaluation, survey research/methodology, and special education, each of whom held a doctorate in his or her content area. The draft survey was revised significantly based on their comments, including cutting the number of questions from 34 to 18 to focus the study and improve the response rate.

The survey instrument and study plan were submitted for approval to the Institutional Review Board and then to the Executive Committee of the American Evaluation Association, which required IRB review as part of its approval process. Following these approvals, the survey was transferred to an electronic SurveyMonkey format for piloting. Pilot participants were recruited through the Southeast Evaluation Association, an AEA affiliate; EVALTALK, the AEA listserv; and personal networks.

The pilot test was conducted to get a sense of non-response rates and variation in responses expected in the full survey (Alreck & Settles, 1995) as well as to test the SurveyMonkey process and survey instrument (Davis, 2005; Rogers, n.d.; Weidenthal, 2002). Specific areas of interest were the approximate length of time for completing the survey and whether there were any problems or enhancements that should be addressed. Pilot participants were asked to complete the SurveyMonkey survey on-line and then complete a one-page protocol (Appendix C) which was based on questions and procedures developed by Davis, (2005), Rogers (n.d.), Simon (2008); Thayer (2006); and Weidenthal (2002).

In total, 29 people stated that they met the survey criteria and volunteered to pilot the survey. They subsequently received an invitation and link to participate in the SurveyMonkey survey. All pilot participants were evaluators with the exception of two people who were asked to test the SurveyMonkey “opt out” feature. Survey responses for the pilot survey are summarized in Appendix D. Response rates for the pilot test were

- Nineteen people responded to the SurveyMonkey survey (65.5%)
- Four people opted out of the survey (13.7%)
- Six people did not respond to the survey invitation (20.6%) of whom two had problems accessing the survey
- Twelve people also responded to the pilot protocol questionnaire (37.9%) by completing the protocol form or sharing comments by telephone or e-mail.

Feedback from respondents who replied to the pilot protocol found no problems with the e-mail invitation, survey design, font, or the time it took to complete the survey, which ranged from “more than 7 minutes” to an outlier high of 30 minutes. Average response time was 15 to 17 minutes. Seven respondents described problems trying to fit their responses into forced choice questions and suggested that the questions be changed and more opportunities be provided to explain their answers. Other comments included adding more information about the study to the survey invitation e-mail, changing the wording of some questions, and adding “don’t know” or “not applicable” to some options. One respondent suggested that the survey include feedback on all the evaluations conducted rather than the most recent one.

Four pilot survey respondents did not receive the original survey invitation e-mail. Three respondents were successfully re-invited with two responding to the survey. One technical problem that blocked the respondent from receiving the invitation e-mail was not corrected. Although survey responses were lost, the pilot process was valuable in learning how to manage the SurveyMonkey process and troubleshoot problems.

Several revisions were made to the survey instrument and invitation e-mail based on pilot participant suggestions for adding more descriptive information to some answer options. Three questions (3, 4, 7) were changed from forced choice to Yes/No responses. The original Likert scale responses (None at all/A slight amount/A small amount/A moderate amount/A great deal) were changed by dropping “A slight amount” and adding “Not applicable/Don’t know” as an option. Question 18 was completely reworded using language suggested by a pilot participant to focus it more on the evaluator-client relationship rather than a general comment on evaluating youth programs.

Data Collection Procedures

The study was conducted via an electronic survey of AEA members located in the United States. SurveyMonkey collected and compiled the raw data which was analyzed by the student researcher. Identity protection for respondents included encryption at the highest available level of the SurveyMonkey data file and the removal of all identifying information from the data files stored on the researcher’s password protected laptop.

Table 4

Calendar of Events for Implementing the Study

Month/Year	Event
January 2008	Approval of dissertation proposal
February/May 2008	Revision of proposal
	Expert panel review of survey instrument
	Revision of survey instrument
May 2008	Submitted study to GWU IRB
July 2008	GWU IRB ruled study as exempt
	Submitted study to AEA Executive Committee
September 2008	AEA approved survey study for dissemination after the AEA annual conference in November
	Selected SurveyMonkey for electronic survey
October/November 2008	Recruited pilot participants and piloted survey
	Revised survey submitted to GWU IRB which found no material changes
	Finalized survey
November 2008	Finalized criteria for selection of AEA member e-mails
December 3, 2008	Initiated survey dissemination and data collection
December 18, 2008	Sent follow up reminders
January 4, 2009	Closed survey collection

Spring 2009	Analyzed survey data
May 2009	Completed survey analysis

The above calendar departed from the original plan in two significant ways. First, the dissemination date for the final survey was delayed due to the need to schedule around other surveys of the AEA membership. December was not the optimal time to disseminate a survey because so many people were out of the office or otherwise occupied during the holiday season, but delaying the survey to February or March 2009 was not an option. Second, although the study was approved for three follow-up e-mail reminders over a period of 6 weeks by the AEA Executive Committee, AEA's schedule permitted only one reminder e-mail and required that the survey be closed 30 days after the original invitation was disseminated.

Sampling Design and Procedures

As noted above, it was not known how many of AEA's approximately 5,500 members had experience working with grantees funded to serve transition-age youth ages 14-25. AEA member data did not include information that would identify those serving the youth age range targeted for the study. AEA also did not allow surveys of the entire membership. Therefore, a request was sent to the AEA Executive Director and Membership Director for e-mail addresses of AEA members who were located in the United States and were members of 12 Topical Interest Groups (TIGs) likely to serve transition-age youth: (a) Alcohol, Drug Abuse, and Mental Health, (b) Assessment in Higher Education, (c) College Access Programs, (d) Crime and Justice, (e) Government Evaluation, (f), Health Evaluation, (g) Human Services Evaluation, (h) Independent

Consulting, (i) Nonprofit and Foundation Evaluation, (j) PreK-12 Education Evaluation, (k) Social Work, and (l) Special Needs Populations.

The resulting list of 3,978 e-mails was inspected. Duplicate addresses and a few e-mail addresses outside the United States were removed. AEA members may join up to five TIGs, but the e-mail list did not identify TIG memberships, so it was not possible to cross-reference TIG memberships to further refine the identification of candidates for the survey invitation. It was hoped that the results of an internal scan of AEA members conducted in 2007-2008 would provide information that would assist in identifying the population of AEA members to be sampled. Unfortunately, information from the scan was not helpful in further focusing the sample to be surveyed, although it did provide additional context for the research study.

The master list of the selected AEA members contained 3,978 e-mail addresses. SurveyMonkey deleted 61 e-mail addresses as “previously opted out.” The researcher deleted 25 e-mail addresses that were duplicates, non-USA addresses, and pilot participants. The resulting list of 3,892 AEA member e-mails comprised the purposive sample which received invitations to participate in the survey. By responding to the survey invitation and completing the survey, AEA members verified that they met the survey criteria.

Two weeks following the initial invitation, the list of respondents was cleaned of opt-outs, bouncebacks, and long term “out-of-office” responses. A second e-mail invitation was sent to the remaining 3,424 respondents who had not yet responded to the survey. Three paper surveys were sent to people who requested them. One respondent explained that she had difficulty reading the survey on the computer screen; the other two

did not explain why they preferred paper surveys—and did not submit completed surveys. The survey closed after 30 days, and the final response tally was

396 (of 443 people who opened the survey) provided useable responses

3,191 people did not respond (other than automatic out-of-office responses)

258 people opted out, had bad e-mail addresses, or were on long-term leave

Instruments and Tools Used or Designed

A structured survey containing both quantitative and qualitative questions was disseminated. The survey contained 18 questions (15 quantitative, three open-ended) and was divided into four sections: Section One: Youth Program Information, Section Two: The Evaluator's Role, Section Three: Planning and Implementation of the Evaluation, and Section Four: Final Section (open-ended questions). Survey guidelines (Alreck & Settle, 1995; SurveyMonkey, 2008) for selecting appropriate question formats, ensuring proper flow, and reducing instrumentation bias were followed in the development of the survey.

Quantitative data was drawn from five forced choice, five yes/no, and five Likert-scale questions. Qualitative data was drawn from three open-ended questions that solicited input on challenges and solutions encountered by the responding evaluator, resources that were helpful or effective in working with youth organizations, and a comment of the respondent's choosing. The invitation e-mail and final survey instrument are found in Appendix A.

Validity and Reliability

Survey questions were based on variables and factors identified in the evaluation literature (see Appendix B). Content and construct validity of the survey instrument were

determined by submitting the draft survey instrument for review to five experts in evaluation and survey research and by pilot testing the instrument with 19 evaluators. Criterion validity was not determined for the survey instrument because no other instrument was found with which to compare it. Internal consistency (reliability) of survey questions with Likert scale responses were calculated by SPSS. Cronbach alphas for the final survey can be found at the end of chapter 4.

Pre-Testing and Piloting Issues

A draft paper survey of qualitative and quantitative descriptive data was developed based on the research questions. The draft survey was sent to five survey/evaluation experts in order to identify survey questions that were unclear or otherwise problematic, to ensure that the questions collected meaningful information and that the content and constructs of the questions were valid, to determine that the survey was a reasonable length, and to verify the need for the study and its potential contribution to the field.

After IRB and AEA Board review, the survey was transferred to an electronic survey format and piloted with 19 evaluators to test the internal consistency of survey questions and the validity of constructs, identify any remaining problems with wording of questions or instructions, and ensure that the electronic survey process proceeded smoothly. The survey instrument was revised after the pilot and before dissemination to the full survey sample.

Data Processing and Procedures of Data Analysis

Data collection and initial processing was provided by SurveyMonkey, an electronic survey provider, supplemented by SPSS and inductive/deductive qualitative

analyses. As noted above, quantitative data analysis included frequencies and percentages for categories of descriptive data. Qualitative, open-ended questions were coded deductively and inductively, i.e., using themes that were identified in the literature as well as new themes that arose from the data.

Before coding qualitative questions, the responses were reviewed to identify patterns or themes. A preliminary set of codes was developed, some of which were suggested in the question itself or in previous questions. The codes were refined via several subsequent passes through the responses. For example, Question 16 asked for information on how challenges to implementation and/or threats to validity were successfully handled during the evaluation. Initial coding was based on the specific types of challenges and threats provided in quantitative questions 13 and 14. Although the question asked for successful handling of an issue, additional coding was added for successful, unsuccessful, on-going, and no handling of the threat or challenge.

Question 17 asked for evaluation resources such as guides, publications, and Web sites that were helpful in working with youth organization evaluations. Initial codes included guides, publications, and Web sites and were expanded to include AEA resources, federal government resources, professional development, evaluator-developed resources, unspecified resources and so on. Question 18 was a comment of the respondent's choosing such as something notable about the evaluator/client relationship, whether the client was representative of other youth organizations served, or an explanation or expansion of previous answers. Initial codes began with those categories, and additional categories were added for comments on the study or survey instrument,

global observations about the evaluator/client relationship or working with youth programs, responses that were not understood, and so forth.

Data verification procedures included examining survey responses for missing data. In total, 47 incomplete surveys were deleted (8 blank surveys and 39 surveys that skipped 7 or more of the 15 quantitative questions). The resulting 396 surveys contained answers to all quantitative questions. SurveyMonkey data was downloaded into SPSS and double-checked for accuracy by visual comparison of data files and comparing frequency summaries with actual answers.

Survey respondents were encouraged to use the open-ended “Other” field to explain previously provided answers to a question or to provide a unique answer that didn’t fit the list of provided choices. Some unique answers were actually subsets of provided answers. In those cases, answers were recoded in forced choice questions or a code was added for the appropriate response category in questions that were not forced choice. This ensured consistency of responses within questions.

Responses to three questions contained a large number of unique “Other” responses that were combinations of the provided choices. A new code and an additional response category were added to the data set as a result. The new category was titled “Combination of some or all of the above” and brought to four the number of survey questions with this choice option. (The fourth question had the “Other” category listed as an existing option when the survey was disseminated.)

Human Participants and Ethics Precautions

Human subjects were protected through compliance with The George Washington University Institutional Review Board procedures and the American Evaluation

Association member survey review process. The study was deemed “exempt” from IRB review due to the low risk of potential harm to study participants. Human subjects were informed of study protections in the survey e-mail invitation and welcome page.

Completing the survey was voluntary, and answers were treated as confidential and reported only in the aggregate. No obvious undue risks were endured, and respondents could stop their participation in the study at any time. Respondents did not benefit directly from participation in the study but might benefit indirectly if the survey results are accepted for publication by the *American Journal of Evaluation*. Care was taken during data collection and analyses to reduce researcher and data handling errors. The researcher doubled checked all data transferred from SurveyMonkey to SPSS visually and by comparing frequencies of both data sets. Data summaries, discussions, and analyses report inconsistencies, anomalies, and unexpected occurrences in the data collection and analysis process in full.

Research Design Limitations

The research design had several limitations. First, it was not known how many of AEA’s approximately 5,500 members worked with grant-funded organizations serving transition-age youth ages 14-25 at the time of the survey. Consequently, it could not be determined if the survey sample of 396 respondents was representative of the population of AEA members or evaluators in general. A second limitation was that responses were based on participant perceptions of the questions, or, as noted by one of the respondents, the questions were “subjective.”

The third, and final, limitation was the low survey response rate of AEA members. There are number of possible explanations based on feedback from pilot and

study participants: lost survey invitations, technical difficulties on the recipient's end such as firewall settings, blocks on SurveyMonkey access, and recipients who did not meet the survey criteria. The timing of the survey, which was disseminated during the December holiday season, was also a factor, especially since 29% of AEA members were employed by colleges and universities (Goodman Research Group, 2008) that customarily close for several weeks during the collection period. Survey fatigue of potential recipients may have been an additional factor affecting response rates given that the study survey was scheduled between two other approved surveys of the AEA membership and followed a major membership scan conducted in 2007-2008.

In summary, the study employed an exploratory, descriptive methodology to collect and analyze a purposeful, but non-representative, sample of American Evaluation Association members who voluntarily responded to a survey composed of quantitative and qualitative questions exploring the factors influencing the evaluator-client relationship.

CHAPTER FOUR

The purpose of the study was to explore factors affecting the evaluator-client relationship in recent evaluations of programs serving transition age youth (ages 14-25) from the perspective of professional evaluators working with the youth grantees and to generate questions about the findings. This chapter presents the results of the quantitative and qualitative data collection and analyses of this study. Simple descriptive statistics were calculated for quantitative questions one through 15. An inductive coding process was used to analyze qualitative questions 16 through 18. Data analyses are organized by the first three research questions guiding the study. The fourth research question will be discussed in Chapter 5.

Research Question One

Type of Youth Organization Evaluated

The following data were collected in response to the forced choice question, “What best describes the type of youth organization you [the evaluator] served?” “Question 1” indicates that this was the first question in the survey.

Table 5

Type of Youth Organization Evaluated

Survey Question 1	Frequency	Valid Percent
Individual non-profit or community-based organization or 501(c)(3)	164	41.4
Public school or school district or component thereof	85	21.5
Four year college or university	33	8.3
Community or junior college	3	0.8
Government agency	31	7.8
Collaboration of all or some of the above groups	64	16.2
Other (please specify)	16	4.0
TOTAL	396	100.0

The category of non-profit/community-based organizations had the highest frequency of youth organizations served by survey respondents (41.4%) followed by public schools or school districts or components thereof (21.5%) and collaborations of types of youth organizations (16.2%). There were 16 unique “Other” responses for this question that included museums, a church organization, a for profit organization, non-public schools, a foundation, a subcontractor to a minority government contractor, research companies, and national youth organizations.

Primary Funding Source for Youth Grant Program Evaluated

The following data were collected in response to the forced choice question, “What was the primary funding source for the grant program that was evaluated?”

Table 6

Funding Source for Youth Grant Program

Survey Question 2	Frequency	Valid Percent
Federal government	204	51.5
State government	69	17.4
Local or municipal government	15	3.8
Foundation	73	18.4
Corporation	4	1.0
Combination of some or all of the above (added later)	20	5.1
Other (please specify)	11	2.8
TOTAL	396	100.0

Over half (51.5%) of the survey respondents indicated that the federal government was the primary funder of the youth program being evaluated, followed by foundations (18.4%) and state government (17.4%). The response category, “Combination of some or all of the above,” was added to this question because of the large number of “Other” responses that described combined funding sources. There were 11 unique “Other” responses for this question that included trust fund, church fund, don’t know, nonprofit funding, cost recovery/fees, and “corporate or foundation fund.”

Primary Youth Target of Grant Program

The following data were collected in response to the Yes/No question, “What type of youth (within the age range 14-25) was the primary target of the grant program?”

Note that “Other” was a comment field and not a response choice.

Table 7

Youth Target

Survey Question 3	Yes Freq.	Yes %	No Freq.	No %	NADK Freq.	NADK %
Academically at-risk youth	235	59.3	129	32.6	32	8.1
Academically gifted youth	35	8.8	302	76.3	59	14.9
Low-income youth	262	66.2	100	25.3	34	8.6
Minority youth	249	62.9	116	29.3	31	7.8
English Language Learners	106	26.8	239	60.4	51	12.9
Youth with disabilities	92	23.2	251	63.4	53	13.4
Gay/lesbian/bisexual/transgender youth	34	8.6	284	71.7	78	19.7
All youth within the geographic scope of the grant	202	51.0	156	39.4	38	9.6
TOTAL 396/100%						
Other (please specify)	92					

A broad range of youth was served by the youth programs being evaluated, with most serving low-income youth (66.2%), minority youth (62.9%), academically at-risk youth (59.3%), and all youth within the geographic scope of the grant (51.0%). Many programs served more than one category of youth. Fewer youth who were academically gifted, gay/lesbian/bisexual/transgender, disabled, or English Language Learners were served by the responding programs, but it is not known whether the numbers reflect a larger trend or a potentially skewed sample. Most “Other” responses were explanations of answers provided to this question. “Other” responses included youth in foster care,

homeless youth, youth at risk of involvement or in the juvenile justice system, pregnant and parenting teens, alcohol and substance abusers, smokers, traumatized youth, physically unfit youth, youth in religious training, teen drivers, and youth interested in nonprofit careers.

Primary Youth Outcome of Grant Program

The following data were collected in response to the Yes/No question, “What type of youth outcome was the primary focus of the grant program?” Note that “Other” was a comment field and not a response choice.

Table 8

Youth Outcome

Survey Question 4	Yes Freq.	Yes %	No Freq.	No %	NADK Freq.	NADK %
High school completion or graduation	186	47.0	180	45.5	30	7.6
Career development or employment	125	31.6	236	59.6	35	8.8
College enrollment or completion	105	26.5	253	63.9	38	9.6
Independent living	53	13.4	300	75.8	43	10.9
Healthy choices	209	52.8	158	39.9	29	7.3
Other (please specify)						
TOTAL 396/100%						
Other (please specify)	164					

Youth programs had a broad range of outcomes with some programs having more than one outcome. The top three categories were healthy choices (52.8%), high school

completion or graduation (47.0%) and college enrollment or completion (26.5%). Some programs had multiple outcomes. This question had 164 “Other” responses, an unusually high number. Most were explanations of answers provided. Unique “Other” responses included religious goals, musical and artistic goals, technology-based learning, service coordination and improvement, and pilot testing tools for schools.

Primary Reason Youth Organization Hired an Evaluator

The following data were collected in response to the forced choice question, “What was the primary reason that the youth organization hired or involved an evaluator or evaluation consultant?”

Table 9

Reason for Hiring an Evaluator

Survey Question 5	Frequency	Valid Percent
Lack of organizational capacity to conduct an evaluation	41	10.4
Funder requirement	254	64.1
External requirement such as accreditation	11	2.8
Stakeholder decision	34	8.6
Board of Director or senior management decision	29	7.3
Combination of some or all of the above (added later)	18	4.5
Other (please specify)	9	2.3
TOTAL	396	100.0

“Funder requirement’ (64.1%) was by far the leading reason for hiring an evaluator or evaluation consultant followed by lack of organizational capacity to conduct

an evaluation (10.4%). “Combination of some or all of the above” was added to this question because of the high number of “Other” responses that described combined funding sources. The nine unique “Other” responses for this question included gathering data, formative evaluation, research and development, credibility, advancing knowledge, don’t know, and “funding, marketing, and decision-making.”

Research Question Two

Type of Organization for Which Evaluator Worked

The following data were collected in response to the forced choice question, “For what type of organization did you work when providing evaluation services to the youth organization?”

Table 10

Type of Evaluator Organization

Question 6	Frequency	Valid Percent
The youth organization being evaluated	41	10.4
Yourself (i.e., self-employed consultant)	68	17.2
Independent consulting firm	83	21.0
Nonprofit services organization/agency	60	15.2
College or university	109	27.5
Government agency	19	4.8
Corporate or private foundation	9	2.3
Other (please specify)	7	1.8
TOTAL	396	100.2

Responding evaluators or consultants worked for a variety of organizations with the largest single employer group being colleges or universities (27.5%) followed by independent consulting firms (21.0%), self-employed consultants (17.2%), and nonprofits (15.2%). Some respondents (10.4%) were internal evaluators who worked for the organization being evaluated. The seven unique “Other” responses for this question included for profit agencies, a medical school, research organizations, college internship, and Americorps/VISTA.

Relationship of Evaluator to Youth Organization

The following data were collected in response to the Yes/No scale question, “What best describes your relationship to the youth organization when providing evaluation services?” Note that “Other” is a comment field and not a response choice.

Table 11

Evaluator Relationship to Youth Organization

Survey Question 7	Yes Freq.	Yes %	No Freq.	No %	NADK Freq.	NADK %
Judge of program merit or worth	208	52.5	178	44.9	10	2.5
Auditor, inspector, or compliance/accountability evaluator	117	29.5	268	67.7	11	2.8
Researcher or scientist	215	54.3	171	43.2	10	2.5
Evaluation facilitator or guide	260	65.7	126	31.8	10	2.5
Team member with evaluation perspective or internal advisor	160	40.4	225	56.8	11	2.8
Collaborator or partner	177	44.7	206	52.0	13	3.3
Program improvement consultant	242	61.1	143	36.1	11	2.8
Empowerment facilitator	75	18.9	304	76.8	17	4.3
Advocate for social change or support of a cause	62	15.7	316	79.8	18	4.5
Educator or capacity builder	171	43.2	211	53.3	14	3.5
Organization development consultant or change agent	115	29.0	263	66.4	18	4.5
Program advocate	67	16.9	310	78.3	19	4.8
Critical friend	104	26.3	274	69.2	18	4.5
TOTAL 396/100%						
Other (please specify)	10					

Many respondents indicated that they had multiple roles as evaluators. Over half of the respondents identified their role as evaluation facilitator or guide (65.7%), program

improvement consultant (61.1%), researcher or scientist (54.3%), and/or judge of program merit or worth (52.5%). All of the “Other” responses for this question were explanations of provided answers with the exception of one unique answer: “technical assistance re survey design and analysis.” Two respondents indicated potentially stressful situations described as (a) “pinch hitter” for the original evaluator who left to focus on a bigger grant and (b) “scope creep” where the evaluator’s responsibilities expanded beyond the original arrangement.

Staff or Stakeholder Anxiety Encountered by Evaluator

The following data were collected in response to the Likert question, “Did you encounter staff and/or stakeholder anxiety related to [the following items]?” Note that “Other” was a comment field and not a response choice.

Table 12

Staff/Stakeholder Anxiety Encountered

Survey Question 8	None at all	A small amount	A moderate amount	A large amount	NA/ Don't know
Lack of experience/knowledge with evaluation	63 15.9%	137 34.6%	126 31.8%	57 14.4%	13 3.3%
Negative past experience with evaluation	145 36.6%	96 24.2%	74 18.7%	32 8.1%	49 12.4%
Fear of failure, negative results, or consequences	70 17.7%	118 29.8%	111 28.0%	81 20.5%	16 4.0%
Organizational culture or social norms resistant to evaluation	141 35.6%	102 25.8%	91 23.0%	45 11.4%	17 4.3%
Role conflicts between evaluator(s) and stakeholders	190 48.0%	114 28.8%	47 11.9%	22 5.6%	23 5.8%
Concerns about evaluator competence (technical or social)	288 72.7%	59 14.9%	22 5.6%	8 2.0%	19 4.8%
Concerns about evaluator understanding of program environment	224 56.6%	103 26.0%	39 9.8%	15 3.8%	15 3.8%
TOTAL 396/100%					
Other (please specify)	13				

Responses to this question indicated that 22.5% to 80.8% of respondents encountered some level of staff or stakeholder anxiety in the categories above. The areas with the highest percentages of anxiety ratings were lack of experience/knowledge with evaluation (80.8%) and fear of failure, negative results or consequences (78.3%). The area with the lowest anxiety ratings was concerns about evaluator competence (22.5%).

The “Other” responses were explanations of answers provided to this question. Three comments indicated that the youth organizations being evaluated were returning clients and were very comfortable with evaluation. Three respondents reported client concerns about how the evaluation results would affect federal funding, whether they would be appropriate for federal reporting requirements, and resistance to receiving news of “anything other than positive effects.” Other concerns included “role conflicts between principal investigators,” and program staff who were distrustful of “outsiders.”

Research Question Three

Primary Focus of Evaluation

The following data were collected in response to the forced choice question, “What was the primary focus of the evaluation?”

Table 13

Focus of Evaluation

Research Question 9	Frequency	Valid Percent
Needs assessment to verify and map the extent of a problem	10	2.5
Formative evaluation to improve and refine the program	58	14.6
Process evaluation to aid program implementation and replication	34	8.6
Outcome/impact evaluation to assess program effectiveness in producing change	172	43.4
Summative evaluation to assess program quality/impact for accountability and decision-making purposes	40	10.1
Combination of some or all of the above (added later)	77	19.4
Other (please specify)	5	1.3
TOTAL	396	99.9

Respondents to this question indicated that outcome/impact evaluation was the most common type of evaluation encountered by respondents (43.4%) followed by a combination of types of evaluation (19.4%) and formative evaluation (14.6%). A new category of response, “Combination of some or all of the above,” was added to this question because of the high number of “other” responses that described multiple evaluation foci. Some respondents explained in their comments that all or most of the options were selected because the evaluation was “comprehensive.” There were five unique “Other” responses that included empowerment or participatory evaluation,

different roles at different phases, evaluation of the relationship between funder and recipient, program review, and “cultural competence/information gathering.”

Evaluation Planning Process Components

The following data were collected in response to the Yes/No question, “Did the evaluation planning process include determinations of [the following items]?” Note that “Other” was a comment field and not a response choice.

Table 14

Evaluation Planning Process Components

Survey Question 10	Yes Freq.	Yes %	No Freq.	No %	NADK Freq.	NADK %
Program evaluability (i.e., determination of program readiness for evaluation)	164	41.4	213	53.8	19	4.8
Agreement/contract terms	276	69.7	99	25.0	21 14	5.3
Evaluation focus	342	86.4	40	10.1		3.5
Methodology (research design)	357	90.2	30	7.6	9	2.3
Evaluation field test (of materials, instruments, processes, etc.)	246	62.1	138	34.8	12	3.0
Data collection	377	95.2	12	3.0	7	1.8
Data editing/reporting	348	87.9	37	9.3	11	2.8
Reporting	365	92.2	23	5.8	8	2.0
TOTAL 396/100%						
Other (please specify)	22	x	x	x	x	

Responses to this question indicated that most (62.1% to 95.2%) evaluation plans included all of the above components with the exception of program evaluability determinations (41.4%). All of the “Other” responses were explanations of answers provided to this question. Two respondents selected “Not applicable/Don’t know” for all their answers and explained that the evaluation was just beginning or that the evaluator came on board after the planning phase. Other comments included “there is little true evaluation planning unfortunately,” “evaluation seems to have been an afterthought,” and one described more attention being focused on research activities than on evaluation processes. Other evaluators described additional or specific plan components such as “cultural, linguistic, and contextual validation,” staff training, stakeholder engagement, logic model development, quality improvement processes, responsibilities, stakeholder presentations, capacity building, and assistance with federal funding requirements.

Youth Involvement in Evaluation Planning and Implementation

The following data were collected in response to the Likert question, “Were youth actively involved in the evaluation planning and implementation process, i.e., in roles other than program participant?” Note that “Other” was a comment field and not a response choice.

Table 15

Youth Involvement in Evaluation

Survey Question 11	None at all	A small amount	A moderate amount	A large amount	NA/Don't know
Youth were actively involved	255 64.4%	80 20.2%	23 5.8%	24 6.1%	14 3.5%
TOTAL 396/100%					

Youth were not actively involved in 64.4% of the evaluations conducted. Youth involvement in evaluation planning and implementation, including youth led evaluation and youth participatory evaluation (YPE), is a relatively new phenomenon, so the low rate of youth involvement was not unexpected. “Other” was not an option for this question, so there were no comments. However a response to Question 18 provides some insight into the pros and cons of actively involving youth in evaluation:

I worked with a multi-site youth program on advocacy for community change. I developed direct relationships with youth to become youth evaluators at each site. Some youth took to it much more than others, so some sites did more youth focused work, but I learned to go with the flow on this. I loved the mentoring and creativity and fun when working with youth, and have tried to carry this same spirit and energy into my other work. I found that the integration with program planning and assessment made me much more a part of the action team than in other projects. We did so many creative activities with youth - photo and interviewing projects, collecting data with palm pilots, community surveys at malls but then immediately analyzed the data and USED IT --wrote press releases and had youth take leadership in sharing the results. So not in any way a "traditional" evaluation role with subjectivity, etc. There were also many challenges. I had to do a lot of the work myself, or work with the youth side by side to ensure more reliability and quality control. It took more time and patience in many ways. But, despite the challenges and extra energy required, the youth empowerment project was my best experience with evaluation to date.

Challenges to Implementation of the Evaluation Plan

The following data were collected in response to the Likert question, “What challenges to the implementation of the evaluation plan perhaps had to be overcome?”

Note that “Other” was a comment field and not a response choice.

Table 16

Challenges to Implementation

Survey Question 12	None at all	A small amount	A moderate amount	A large amount	NA/Don't know
Rigid organizational procedures or settings	153 38.6%	131 33.1%	75 18.9%	23 5.8%	14 3.5%
Inadequate/Problematic data	60 15.2%	117 29.5%	112 28.3%	95 24.0%	12 3.0%
Disagreement/Debate over who should conduct the evaluation	323 81.6%	38 9.6%	14 3.5%	7 1.8%	14 3.5%
Lack of staff cooperation, collaboration, or communication	152 38.4%	146 36.9%	65 16.4%	27 6.8%	6 1.5%
Limited methodological capacity of staff	165 41.7%	112 28.3%	67 16.9%	35 8.8%	17 4.3%
Lack of staff resources such as time or funding	118 29.8%	123 31.1%	98 24.7%	50 12.6%	7 1.8%
Unexpected/Unanticipated evaluation results or consequence	184 46.5%	136 34.3%	40 10.1%	13 3.3%	23 5.8%
Friction between youth organization and funder	276 69.7%	67 16.9%	19 4.8%	5 1.3%	29 7.3%
Misapplication of theories, e.g., constructionism or post-modernism	308 77.8%	24 6.1%	8 2.0%	5 1.3%	51 12.9%
TOTAL 396/100%					
Other (please specify) = 11					

Over half of the respondents reported challenges related to inadequate or problematic data (81.8%); lack of staff resources such as time or funding (68.4%); lack of staff cooperation, collaboration, or communication (60.1%); rigid organizational procedures or settings (57.8%); or limited methodological capacity of staff (54.0%). 47.7% of respondents reported challenges related to unexpected or unanticipated evaluation results or consequences. All “Other” responses were explanations of answers provided to this question. Most described funder requirements that negatively impacted evaluation such as short time frames, shifting evaluation priorities, burdensome data and/or reporting mandates, evaluation requirements that were inappropriate for the culture/context of the evaluated program, and unnecessary requirements. Other issues included lack of reliable, valid assessment tools; multiple stakeholders; and difficulty in engaging youth in the program to a lesser degree.

Threats to Validity of the Evaluation

The following data were collected in response to the Likert question, “What threats to validity of the evaluation plan perhaps were encountered? Note that “Other” was a comment field and not a response choice.

Table 17

Threats to Validity

Survey Question 13	None at all	A small amount	A moderate amount	A large amount	NA/Don't know
Outside forces affected intervention or results	116 29.3%	126 31.8%	73 18.4%	40 10.1%	41 10.4%
Interventions mutated from the original design	114 28.8%	107 27.0%	95 24.0%	43 10.9%	37 9.3%
Youth reacted to please or displease program operator (halo/cloud effects)	205 51.8%	71 17.9%	27 6.8%	2 0.5%	91 23%
Youth became sensitized to intervention measures over time	206 52.0%	66 16.7%	19 4.8%	4 1.0%	101 25.5%
Maturation of youth participants biased the observed results	202 51.0%	71 17.9%	27 6.8%	7 1.8%	89 22.5%
Evaluation process caused disruptions that impacted results	245 61.9%	84 21.2%	19 4.8%	1 0.3%	47 11.9%
Measurements became unstable or changed over time	224 56.6%	73 18.4%	43 10.9%	11 2.8%	45 11.4%
Control groups were not comparable (initially or over time)	77 19.4%	52 13.1%	38 9.6%	35 8.8%	194 49.0%
Program interventions migrated to the control groups	144 36.4%	34 8.6%	17 4.3%	4 1.0%	197 49.7%
TOTAL 396/100%					
Other (please specify) = 22					

The majority of respondents reported experiencing threats to validity related to mutating interventions (61.9%) and outside forces (60.3%). It is also interesting to note the low rate (6.3%) of respondents reporting disruptions caused by the evaluation

process. This question had much higher rates of Not applicable/Don't know responses than other questions especially for the items related to control groups (49.0% and 49.7%) as well as for youth "halo" or "cloud" effect reactions (23.0%), youth sensitization to interventions (25.5%), and youth maturation (22.5%). All of the "Other" responses were explanations of Likert scale responses and provided possible explanations for the high number of NA/DK answers. Six respondents said their programs had no control groups; five said it was too early in program or evaluation implementation to answer the questions. Other threats described were difficulties with obtaining youth participation, youth attrition, data problems, and required academic assessments that were modified during the course of the evaluation.

Evidence of Successful Evaluation Implementation and Usefulness

The following data were collected in response to the Likert question, "What evidence of successful evaluation implementation and usefulness perhaps were observed?" Note that "Other" was a comment field and not a response choice.

Table 18

Evidence of Success

Survey Question 14	None at all	A small amount	A moderate amount	A large amount	NA/Don't know
Improved stakeholder focus, ownership, and/or skills	16 4.0%	93 23.5%	145 36.6%	94 23.7%	48 12.1%
Improved organizational capacity, impact, viability, and/or credibility	16 4.0%	82 20.7%	163 41.2%	86 21.7%	49 12.4%
Improved program implementation	18 4.5%	77 19.4%	157 39.6%	94 23.7%	50 12.6%
Improved program results	20 5.1%	99 25.0%	130 32.8%	54 13.6%	93 23.5%
Improved use of evaluation results by youth program, funder, or other audience	22 5.6%	82 20.7%	127 32.1%	98 24.7%	67 16.9%
TOTAL 396/100%					
Other (please specify) = 22					

The vast majority, 71.4% to 83.6%, of respondents reported evidence of successful evaluation implementation and usefulness. All “Other” responses were explanations of Likert scales answers provided to this question. Eleven respondents reported that evaluation results were still under development or not reported yet. Additional evidence of success included design changes, program modifications, improved understanding of barriers to strong program implementation, program

expansion, and additional funding. A few “Other” responses described a lack of success or sharing results:

1. When I came on board, I informed the client and researchers that the instrument used would not measure the program appropriately...I introduced some methods...but it was too late in the game to be of real value
2. Program pulled prior to formal release of evaluation results due to budget issues.
3. Most times I have no idea if the district even read what I wrote--or cared. The point was to fulfill an obligation, but to me, it was to positively impact program implementation
4. I felt that the organization only hired [us] to meet the requirements of the funder and did not use our findings to make any program improvements. Our report was passed along to the funder and, I believe, that is where it ended.

Audiences with Whom Evaluation Results were Shared

The following data were collected in response to the Likert question, “With whom were evaluation results shared?” Note that “Other” was a comment field and not a response choice.

Table 19

Audiences for Evaluation Results

Survey Question 15	Yes Freq.	Yes %	No Freq.	No %	NADK Freq.	NADK %
Program sponsors and/or funders	369	93.2	8	2.0	19	4.8
Program director and staff	380	96.0	3	0.8	13	3.3
Program participants	152	38.4	164	41.4	80	20.2
Practitioners in the youth service field	196	49.5	123	31.1	77	19.4
Policymakers and/or opinion leaders	212	53.5	109	27.5	75	18.9
Social scientists	117	29.5	203	51.3	76	19.2
Other evaluators	186	47.0	144	36.4	66	16.7
The general public	119	30.1	187	47.2	90	22.7
TOTAL 396/100%						
Other (please specify) = 36						

Respondents reported that almost all programs shared evaluation results with program directors and staff (96.0%) and program sponsors or funders (93.2%). About half shared responses with policymakers/opinion leaders (53.5%), practitioners in the youth service field (49.5%), and other evaluators (47.0%). All “Other” responses were explanations of answers provided to this question. Several respondents said it was too early to share results yet; of these some responded NA/DK or answered for the groups with whom they intended to share. Some respondents explained that their answers were a mix of actual and planned recipients. Others indicated that the evaluation results were

available on web sites of the youth organization, project, and/or funder. Additional audiences were identified including peer reviewed journals, prospective funders, other foundations, boards of directors, volunteers, interested community stakeholders or leaders, project advisory groups, school administrators, principals, resource teachers, university leadership, the public health and tobacco community, and “anyone.”

Reported problems with sharing results included:

1. No results shared due to program termination
2. I don't think results went anywhere, but don't know.
3. The results were regarded as unfavorable toward the program, shared only internally and not subsequently published or given to the school board.

Qualitative Data Analysis

SurveyMonkey data indicated that 363 people responded to each of the survey's remaining three questions, which were open-ended and qualitative. However, because the survey required an answer for each question, many of the “responses” were actually non-responses, as described below. Consequently, the number of evaluators responding to the qualitative survey questions ranged from 188 for Question 18 to 267 for Question 16. It is possible that some respondents who provided detailed comments for the quantitative questions may have decided not to repeat their answers in the qualitative questions.

Implementation Challenges or Threats

Survey question 16 asked respondents to “Please describe a serious implementation challenge or threat to validity that you encountered and successfully dealt with during the afore-mentioned evaluation. What was the challenge or threat, how did you deal with it, and with what effect?”

Some types of implementation challenges or threats to validity were reported by 267 respondents of whom 31 indicated that no responses or threats were encountered. The rest were non-responses such as “NA,” “Not applicable,” “xxx”, etc. A summary of the issues described in the responses is provided in the following table. Note that some respondents described more than one type of challenge or threat.

Table 20

Resolution of Implementation Challenges or Threats

Survey Question 16	Frequency
Data/Management Information Systems	62
Staff/agency/school cooperation, collaboration, communication, trust or comfort	46
Fidelity to evaluation plan, intervention or methodology	46
Youth recruitment, participation, retention, or response rates	38
Measurement (unstable, difficult to measure, biased)	28
Methodology	27
Control/comparison groups, random assignment, sampling	22
Funder expectations, communications, changes, reporting	21
Leadership/staff turnover, resources, time, funding	20
Outside forces	18
Program resources, capacity or expertise for conducting evaluation	18
Cultural or diversity issues	8
Unexpected/unanticipated evaluation results or consequences including the impact of the evaluation on the program	5
Other	15
Total	374

Although the question asked respondents to describe challenges or threats that were deal with successfully, only 58 respondents specifically stated that the issues were successfully resolved. Some respondents described the challenge and the actions taken to

address them but did not state an outcome or whether it achieved the desired result.

Other respondents indicated that the problem was on-going or had not yet been resolved or merely stated their problems without describing the actions taken to address them. A sample of responses follows.

The state feared a lawsuit if I found and wrote negative things. The state did not collect routine data to evaluate the program, and they would not let me collect important types of data. I evaluated the parts that they would let me, and then I let them conduct their own evaluation of the parts for which they wouldn't share data with me. I provided feedback to them on their separate evaluation along the way. Most deficits were elided over by their internal reporter.

Resistance to evaluation on the part of facilitators, who felt that evaluation was intrusive. We surveyed participants and asked about their reactions to evaluation, and they were strongly positive about evaluation. This evidence appears to be somewhat convincing to at least some facilitators.

Pragmatic scheduling around various changes in school schedules and teacher avoidance for observations/interviews. Developing a trusting relationship (which did not occur until 1.5 years into a 5 year grant) where evaluators and researchers are personable and appear interested in the participants' wellbeing. This generated a more responsive participant pool who were more likely to work with evaluators/researchers to provide substantive data.

Challenge: Gathering data from dispersed groups of former participants. Solution: Random sampling of former participant group. Result: Representative sample of group with acceptable error rates.

As always there is an initial reluctance from the staff to an outside evaluator. We work to establish rapport as time goes on, which enables us to provide an effective evaluation of the program.

Challenge – enrolling youth and their parents into a study re mental health. Strategy – a work group of parents of children's mental health consumers, clinicians, and evaluators brainstormed effective methods of recruitment (e.g., clinicians would introduce study & seek consent to be called by evaluator; evaluator & clinicians would remain in contact re families availability). Effect – we greatly increased our recruitment rates and improved relationships among our division, clinical staff, & families.

Resources for Youth Organization Evaluations

Survey question 17 asked respondents, “What evaluation resources (guides, publications, Web sites, etc.) do you find to be particularly helpful or effective in working with youth organizations (not just the youth program described previously)?”

Of the 363 “responses” to this question, 262 described one or more resources and/or made a comment. A summary of the types of resources identified is contained in the following table. The complete list of resources can be found in Appendix E.

Table 21

Evaluator Resources Summary

Survey Question 17	Frequency
AEA Resources	53
The Joint Committee on Standards for Education Evaluation	8
Texts/Reference Books	27
Journals/Research	14
Other Publications	35
Government Resources	51
Foundation Resources	16
Websites/Internet Resources	41
Peer/Stakeholder Networks/Communications	25
Instruments/Data/Software/Technology	23
Experience/Training/Professional Development/Education	15
Evaluator/Stakeholder Developed Resources/Procedures	10
Other Resources	43
Nonspecific Resources	77
None/Don't Know/Can't Recall	48
Non-Responses	58
Unclear Response (NREP, ISE recommendations, AOI)	3
Comments not Related to Resources	4
Total Responses	551

Many respondents indicated more than one type of resource used; their answers are reflected in more than one category. The Nonspecific Resources category contained responses such as “websites” or “texts” or “anything to do with participatory evaluation.” Non-Responses contained responses such as “NA,” “Decline to comment,” “X,” and the like. As expected, a large number of AEA and federal government resources were identified, given that respondents were AEA members and over half identified the federal government as the primary funding source.

Specific resources that were mentioned repeatedly included books by Michael Quinn Patton, particularly *Utilization-Focused Evaluation*; W. K. Kellogg Foundation evaluation resources; Harvard University resources particularly the Harvard Family Research Project; The Joint Committee on Standards for Education Evaluation, and Western Michigan University’s Evaluation Center. There were several unexpected categories that indicate creativity of the part of evaluators and a possible move away from more traditional sources of information, or ways of finding information, such as publications, journals, and Web sites. These categories were Peer/Stakeholder Networks/Communications, Instruments/Data/Software/Technology, and Resources/Procedures Created by Evaluators/Stakeholders.

Many respondents made a distinction between resources related to evaluation and those related to youth programs. The large number of “Don’t Know” responses most likely refers to the latter. Two of the four comments did not relate to resources and are included elsewhere in this document. The two others were illuminating:

I was only an advisor in this project and not directly involved with using resources.

Sometimes I pull out AEA's journal, but it feels like my work--local, often an afterthought--is so far removed from the academic presentation of evaluations. Control groups?! HA! All I can do is bring my best training to the table, and provide the most appropriate evaluation for their needs given the situation. It is not ideal, but it is reality.

Open Comment Responses

Final survey question 18 asked respondents to “Use this question to make any comments of your choosing. For example, was there something notable about the evaluator/client relationship that was not captured in the previous survey questions? Was the client described representative of other youth organizations you have served? Would you like to explain or expand upon any of your previous answers?”

Table 22

Evaluator Comments

Survey Question 18	Frequency
Something notable about the evaluator/client relationship	10
Client representative of other youth organizations served	6
Client NOT representative of other youth organizations served	14
Explain/Expand on previous answers	101
General observation about evaluation/evaluation of youth programs	41
Comment on study or survey instrument	13
Answer/Part of answer not understood	7
Could not respond to all survey questions because evaluation described was not complete or too early in implementation	6
Respondent not sure he or she was the “right person” to take the survey or that his/her answers should be included	4
Answers referred to more than one youth program	3
Wanted information on youth evaluation resources, specifically qualitative methods such as focus groups and interviews	1
*Total Responses	206
Indicated no response	175
Total Respondents	363

*Some respondents had multi-part answers that were included in more than one category above.

Of those who responded to this question, 49.0% explained or expanded on survey answers, and 19.9% included a general observation of some type. Four people expressed

concern that they were not the right person to respond to the survey, and three said there answers referred to more than one youth program evaluation. Of the comments that described positive or negative evaluator/client relationships or evaluation experiences, many indicated positive, successful relationships or experiences with clients. However, more responses described challenges to the relationship or to evaluation implementation. This apparent negative skew should be taken in context since many people did not respond to the question and, in a previous question, the vast majority of respondents reported evidence of successful evaluation implementation and/or usefulness.

A sampling of comments include:

In my experience, client relationship building is critical to an effective evaluation. Open lines of communication that can be initiated by either party are essential. While face to face kick-off meetings help get a study off the ground, constant coordination, communication, instruction, counseling, and encouragement are needed to sustain the effort, obtain necessary data, and ensure that questions of interest to the client are being answered by the research.

The client-evaluator relationship was uniquely challenging in that all the client stakeholders had a qualitative orientation and a low level experience in social science research methods, while our evaluation team was primarily quantitative. The client's perspectives were not adequately understood by/expressed to our team prior to accepting the evaluation contract. The differences in perspectives posed significant challenges to implementing our evaluation plan and using the results.

Cultural understanding and Spanish language capacity were key to developing a trusting relationship.

Evaluation is a term that overwhelms many nonprofit clients... above all, clients need help making evaluation tools apply to their situation. While I may use a theory-guided approach, I find that the jargon of evaluation theory to be confusing and inefficient in building solid relationships with clients. Listening is among the top skills that good evaluators must have... and, evaluators have to be willing to do the heavy lifting, including data entry to get projects rolling... and show clients the value of analyzed data.

I am always amazed that grantees do not consult with the evaluator during the grant writing process. It always seems to occur at the very end. Grantees do not

seem to see the connection between grant program planning and the development of an evaluation plan. They should go hand and hand, but it seems more often as an add-on to the grantee.

I think Fidelity of implementation is a huge issue. The other big challenge is the focus (from the Feds) on outcome measures that would be unrealistic to expect in the given time frame.

Much of my work is organizational development occasioned by the need for service agencies to respond to formal mandates that exceed current infrastructure capacity or development formal [sic]. These agencies face the very real possibility of being forced to mutate their programs in undesirable ways to accommodate the external demands and are seeking ways to make the best response in terms of intervention integrity v meeting imposed requirements.

There is a dimension you aren't capturing here that being the press for social desirability of the contracting agent who wants to retain state project funding, thus impacting how closely they want to be involved and shape the evaluation and related findings, similarly then the press to have the evaluation be for youth and toward understanding whether the programs were doing "good things" were muted and created tensions. ... I could write a whole paper on this challenge to the client relationship.

We have a truly collaborative partnership that has been extremely successful.

The evaluation of this 3 year grant went very smoothly with all parties engaged and pleased with the process and the results.

...Being an "internal" evaluator often has its pros (easier access to data; more frequent contact & communication with program; not being seen as an outsider; more familiarity with the context & demands being handled by the program staff), but it also has its cons (e.g., having dual responsibilities -- if the program manager is aware of but does not share less than favorable evaluation results with the district leadership, what then becomes the internal evaluator's duty?)

The following comment sums up the importance of program evaluation and the evaluator/client relationship rather nicely:

What is most notable is how satisfying a highly participatory evaluation is for everyone involved. We all want a compelling sense of the future and try to accomplish that. Evaluation is a solid gateway for organizations and individuals to get there.

Instrument Reliability

Internal consistency/reliability was calculated for the five-point Likert scale questions using SPSS. The resulting Cronbach alphas are reported in Table 23 and met or exceeded the 0.70 threshold identified by Fraenken and Wallen (1996, p. 163) for most research. McMilland and Schumacher (1993, p. 231) noted that reliability coefficients as low as 0.50 may be acceptable for exploratory research.

Table 23

Internal Consistency/Reliability of Instrument

Question Number	Cronbach Alpha
Question 8:	0.802
Question 11:	Unable to calculate (single scale)
Question 12:	0.730
Question 13:	0.698
Question 15:	0.863
Questions 8, 11, 12, 13, 15 combined	0.831

Summary of Results

A purposeful sample of members of the American Evaluation Association was invited to participate in a survey on the relationship of evaluators and grantees serving youth in the age range 14-25 years. Three hundred and ninety-six members responded. A broad range of responses was received for most questions.

Over half of the survey respondents indicated that the federal government was the primary funder of the youth program being evaluated, with most serving at-risk youth populations. Many programs had multiple youth outcomes. “Funder requirement” was the most common reason cited for hiring an evaluator or evaluation consultant.

Responding evaluators or consultants worked for a variety of organizations and had multiple roles as evaluators with the most common being evaluation facilitator or guide, program improvement consultant, researcher or scientist, and/or judge of program merit or worth. Most respondents reported staff or stakeholder anxiety related to the evaluation. Outcome/impact evaluation was the most common type of evaluation.

Most evaluation plans included seven components: contract terms, evaluation focus, methodology, field tests, data collection, data editing or reporting, and reporting of results. Fewer than half of respondents reported that determination of program readiness for evaluation was a component of the plans. Almost two thirds of respondents reported that youth were not actively involved in evaluation planning or implementation.

A majority of respondents reported evaluation challenges related to inadequate or problematic data; lack of staff resources such as time or funding; lack of staff cooperation, collaboration, or communication; rigid organizational procedures or settings; or limited methodological capacity of staff. Just under half of respondents reported challenges related to unexpected or unanticipated evaluation results or consequences. A majority of respondent reported threats to validity related to mutating interventions and outside forces. On a more positive note, very few respondents reported disruptions of their programs caused by the evaluation process, and the vast majority of respondents reported evidence of successful evaluation implementation and usefulness in five

categories related to stakeholders, organizational effectiveness, program implementation, program results, and use of evaluation results. Respondents also reported that almost all programs shared evaluation results with program directors/staff and program sponsors/funders. About half shared results with policymakers/opinion leaders, practitioners in the youth service field, and other evaluators.

When asked to describe challenges or threats that were dealt with successfully, few respondents specifically stated that the issues were successfully resolved. Some respondents described the challenge and the actions taken to address them but did not state an outcome or whether it achieved the desired result. Others indicated that the problem was on-going or had not yet been resolved or merely stated problems without describing the actions taken to address them.

Resources used by evaluators spanned a wide range of content and delivery methods. A few respondents indicated that no resources were used or identified. The complete list of resources can be found in Appendix E.

The last survey question asked respondents to make comments of their choosing. Almost half of the survey respondents did not respond to the question. Of the comments that described evaluator/client relationships or evaluation experiences, many indicated positive, successful relationships or experiences with clients. However, they were outnumbered by responses describing challenges or problems with the relationship or evaluation implementation.

The apparent negative skew in responses to the qualitative question on challenges or threats and the open comment question should be taken in context. Many people did

not answer these questions, and, in a previous question, the vast majority of respondents reported evidence of successful evaluation implementation and/or usefulness.

Conclusion

Because this was an exploratory study and it was not known if the survey sample was representative, the results of the survey cannot be generalized and should be interpreted with caution. Survey data described in this chapter indicate wide variation in the situations and experiences described by evaluators. What the data may mean and possible implications for research, policy, practice, and professional development will be discussed in the next chapter.

CHAPTER FIVE

Programs and systems serving youth ages 14-25 in the United States operate in a climate where competition for grant funding is fierce, where there is an increasing demand for results, and where the penalties for poor results may be severe. Funding, credibility, and viability of youth serving organizations may be significantly impacted by monitoring and evaluation results. The conceptual framework for the study is based on constructivist, organizational change, and evaluation theories in which relationships are critical components of successful endeavors. Consequently, the research questions for this exploratory, descriptive study sought to investigate the characteristics of youth programs, evaluators, and evaluation planning and implementation and to generate questions related to the evaluator-client relationship based on the findings.

Summary of Findings

Research Question One: Characteristics of Youth Organizations

Not surprisingly, a broad range of youth service organizations were served by evaluators responding to the survey including non-profit and community-based organizations, public schools, and collaborations of organizations from the private, public, and nonprofit sectors. Interestingly, slightly more than half of the survey respondents indicated that the federal government was the primary funder of the youth program being evaluated, followed by foundations and state government.

Most programs served low-income youth, minority youth academically at-risk youth, and all youth within the geographic scope of the grant with many programs serving multiple categories of youth and multiple youth outcomes. The most common

youth outcomes were healthy choices, high school completion or graduation, and college enrollment or completion. The most common reason for hiring an evaluator or evaluation—by far—was funder requirement followed by lack of organizational capacity to conduct an evaluation.

Research Question Two: Characteristics of Evaluators

Responding evaluators or consultants worked for a variety of organizations with the largest single employer group being colleges or universities followed by independent consulting firms, self-employed consultants, and nonprofits. Some respondents were internal evaluators who worked for the organization being evaluated. Most respondents indicated that they had multiple roles as evaluators, with some noting that their role changed or evolved during the course of the evaluation. Over half of the respondents identified their roles as evaluation facilitator or guide, program improvement consultant, researcher or scientist, and/or judge of program merit or worth.

A large number of respondents reported staff or stakeholder anxiety in seven categories of concerns, experiences, or fears. The highest percentages of anxiety ratings were due to (a) lack of experience/knowledge with evaluation and (b) fear of failure, negative results or consequences. Interestingly, the area with the lowest anxiety ratings was concerns about evaluator competence.

Research Question Three: Evaluation Planning and Implementation Characteristics

A number of factors identified in the literature as characteristics of, challenges to, or threats to evaluation implementation were presented based on the assumption that they would be likely factors to influence the evaluator-client relationship. Outcome/impact evaluation was the most common type of evaluation in which respondents were involved,

followed by evaluations with multiple foci, and formative evaluation. Most evaluation plans contained seven components (contract terms, evaluation focus, methodology, field tests, data collection, data editing or reporting, and reporting of results). Fewer than half of respondents included an eighth component, determination of program readiness for evaluation.

Youth involvement is a relatively new, but growing, aspect of youth program evaluation, so it was not surprisingly that most respondents reported that youth were not actively involved in evaluation planning and implementation processes. Over half of the respondents reported evaluation challenges related to inadequate or problematic data; lack of staff resources such as time or funding; lack of staff cooperation, collaboration, or communication; rigid organizational procedures or settings; or limited methodological capacity of staff. Respondents also reported challenges related to unexpected or unanticipated evaluation results or consequences. Threats to validity related to mutating interventions and outside forces were reported by a majority of respondents. The question about threats to validity had much higher rates of Not applicable/Don't know responses than other questions especially for the items related to control groups and youth responses to program interventions. These responses may indicate that some programs did not use control groups and/or that it was too early to determine youth responses to the intervention.

On a more positive note, the majority of respondents reported evidence of successful evaluation implementation and usefulness in five categories related to stakeholder ownership and skills, organizational effectiveness, program implementation, program results, and use of evaluation results. Respondents also reported that almost all

programs shared evaluation results with program directors/staff and program sponsors/funders. About half shared results with policymakers/opinion leaders, practitioners in the youth service field, and other evaluators.

Responses to open-ended questions contained more descriptions of evaluation problems than successful resolutions to them, including specific descriptions of the evaluator-client relationship. However, many survey participants did not answer these questions, and, in a previous question, the majority reported evidence of successful evaluation implementation and/or usefulness on multiple fronts. Therefore, the apparent negative skew in responses to the qualitative question on challenges or threats and the open comment question should be taken in context. Finally, the survey suggests that the evaluator-client is enhanced by the use of a number of resources across a wide range of content and delivery methods; although a few respondents indicated that no resources were used or identified. The complete list of resources can be found in Appendix E.

Questions Generated by Findings

The findings describe both positive and negative aspects of evaluator relationships with their youth grantee clients. It does appear that the characteristics or factors identified in the literature and explored in the survey questions may have an influence on the evaluator-client relationship, especially as described in the qualitative responses. More information on this important relationship is needed.

Survey responses also supported the study's theoretical foundation and conceptual framework. Survey respondents described both cooperation with and resistance to evaluation processes on the individual and organizational levels as well as the importance of interpersonal relationships in accomplishing the work of the evaluation. They also

described the different perspectives on and reactions to the evaluation process on the part of program staff and stakeholders. It appears that the researcher-constructed concept of “parallel universes” where one universe is inhabited by program staff and stakeholders and the other by evaluators is valid and that internal and external forces do influence the evaluator-client relationship—positively by moving the parallel universes closer together or negatively by pushing them apart—thereby impacting the effectiveness of the evaluation.

These findings generated a number of issues and questions which have serious implications for funders, evaluators, evaluator training programs, and youth grantees. Youth grantees served multiple, hard-to-serve, at-risk youth populations via programs and interventions that had goals of improving outcomes (high school completion or graduation, healthy choices, independent living, and college enrollment or completion) that have stubbornly resisted improvement over the years. It is telling that the reason most grantees involved an evaluator is that the funder required them to do so and that they subsequently struggled with several aspects of evaluation according to the evaluators with whom they worked.

Evaluators reported their own struggles when working with program staff and stakeholders. Responses to quantitative questions described varying levels of client anxiety and cooperation, problems with data collection, rigid organizational structures, and lack of resources to support the evaluation. Evaluators reported a variety of evaluator roles, some of which changed over time as the evaluation proceeded, ranging from judge of merit or worth to advocate for social change. Open-ended survey responses indicated varying degrees of direct involvement with clients and program services, familiarity with

program interventions, and personal contact with participating youth. The latter could be one explanation for the relatively large numbers of “Not applicable/Don’t know” responses to the questions about threats to validity based on youth reactions to the interventions.

The responses to the open-ended questions, in particular, provided insight into challenges faced by evaluators in implementing the evaluation and establishing a positive working relationship with the client. The depth of client concerns about evaluation results and how they would be used, the disruptions caused when funders changed the rules during the course of the evaluation, the lack of resources to adequately fund the evaluation process, the difficulties in establishing trust and rapport with the client, and the satisfaction (or frustration) when everything clicked into place (or not) were very real factors that influenced the evaluator-client relationship. The fact that only some of these challenges were successfully addressed is another indication that more research and information is needed on the evaluator-client relationship.

A sampling of the questions generated by the study findings include

- Will additional studies support the construct of “parallel universes” inhabited by evaluators and clients? Will strategies for building trust in the evaluator-client relationship emerge?
- Does the amount or percentage of grant funds make a difference in the commitment of staff to program evaluation, evaluator quality, and final evaluation outcomes?
- Is evaluation really an “afterthought” as described by some survey respondents? If so, is there a way that evaluation planning can be made part of program planning

from the beginning rather than adding it on when grant applications are submitted?

- What is the relationship between client anxiety levels and factors such as type of evaluator role, evaluation focus, challenges to implementation, or threats to validity? For example, are more formal evaluator roles such as judge of merit or worth correlated with higher levels of client anxiety than less formal roles such as collaborator or team member?
- How should the finding that youth program clients indicated a low level of concern about evaluator competence be interpreted, particularly since it may be difficult for evaluators to detect or interpret the concern?
- Are evaluator roles correlated with external variables such as the foci of the evaluation or internal factors such as the evaluator's personality or training?
- How do evaluator roles impact the evaluator-client relationship? Do some roles, such as judge or auditor, have a more negative impact on the relationship? Do others, such as critical friend or collaborator, have a more positive impact?
- Does evaluator training or credentialing influence client anxiety or resistance?
- What is the relationship between evaluation foci and anxiety levels? For example, are impact evaluations that assess program effectiveness and summative evaluations that assess program quality for accountability purposes correlated with higher levels of client anxiety than formative and process evaluations which have goals of improving the program or aiding program implementation or replication?

- What types of external forces impact evaluations? What strategies can be developed to address them?
- How can evaluators and clients address fidelity problems with both program and evaluation methodologies?
- What strategies can be used to recruit and retain youth program participants?
- What additional populations (youth organization directors, funders, youth, parents, program staff...) should be surveyed to provide a fuller picture of the evaluator-client relationship?
- What is the impact of current environmental factors (increased competition for funding, increased funder demand for positive results) on the evaluator-client relationship?
- Does the evaluator-client relationship differ based on the type of program or provider, e.g., is the relationship different for evaluators working with youth service providers versus adult service providers, with programs serving youth with disabilities versus those serving youth without disabilities?
- What other data collection methods could be used to collect data? Would focus groups or case studies provide richer data on the evaluator-client relationship?

Recommendations for Further Research

The questions generated by the study findings suggest several avenues for further research, such as refinement and validation of the survey instrument, complementary surveys, and in-depth investigations, which are described below.

Refinement and Validation of the Survey Instrument

Analysis of survey results and feedback from survey participants indicated several aspects of the survey process that could be improved. An obvious improvement would be to deploy the survey at a more auspicious time of year; the winter holiday season was not an optimal time for collecting data.

The type of question or question format needs to be revisited for some questions. A few respondents said they were confused about questions asking for primary youth target and primary outcome because multiple Yes or No responses were permitted. Other respondents disliked having to respond to each question and preferred to select just the answers that applied to them. One mentioned that having to respond to each question was frustrating and might cause people to give up before completing the survey. In fact, eight people opened the survey but did not respond to it, and 39 people completed one or two sections of the survey before stopping. (It should be noted that several question formats were changed from forced choice to Yes or No on the basis of feedback from pilot participants.)

Another formatting issue that needs to be resolved is whether to separate “Not applicable/Don’t know” responses into two separate responses in order to gain more detailed information. For example, as one respondent pointed out, the responses to sub-questions on threats to validity related to control groups could be separated into Don’t Know and Not Applicable, which would give a better indication of evaluations that did not use methodologies using control groups. For the purposes of this exploratory study, which was focused on the “yes” responses, it was deemed unnecessary to go to that level of detail.

The issue of whether to require an answer to each question and sub-question should also be examined. A few respondents commented that they found having to respond to each question burdensome. They would have preferred to select only the answers that applied to them and to skip the questions that they did not want to answer. The decision to require an answer was made by the researcher to ensure that respondents answered each question completely and thereby provide a fuller picture of factors impacting the evaluator/client relationship. Trade-offs such as reducing the large number of “non-responses” generated by the open-ended questions versus ensuring that people did not inadvertently skip questions would have to be considered.

The terminology of some questions should also be revisited to ensure that terms commonly used by youth service practitioners—but not necessarily known to evaluators—are defined. For example, “healthy choices” include those related to physical health, mental health, nutrition, alcohol and drug use, youth violence, sexual activity, suicide prevention, illegal activities, and so on. Lack of familiarity with these terms may explain why some evaluators listed subsets of healthy choices as unique “Other” responses rather than simply selecting the “healthy choices” option. A couple of respondents said the terms, “results” and “client,” were not defined. One suggested “outcomes” rather than “results,” and the other wanted to know if the “client” was (a) the agency or organization that hired the evaluator or (b) the youth themselves.

Complementary Surveys

One respondent noted that survey questions were “very subjective” which highlights the fact that survey responses reflect the evaluator view of the evaluator-client relationship. A survey of youth services operators would provide information from the

client's perspective and would be a valuable addition to the literature. Convincing youth services operators to respond to the survey might be a challenge given their discomfort with evaluation described in the current survey. Surveys of additional stakeholders such as funders, youth, parents, and youth program staff would be beneficial and would provide a 360 degree view of the evaluation process and the evaluator-client relationship.

The survey could also be reworked and deployed to explore other types of program evaluations (such as those serving adults over the age of 25) and select cases such as federally-funded youth programs or those serving youth with disabilities. Common concerns, strategies, strengths, and weaknesses that span all types of evaluations or clients, as well as unique factors found only in the subsets, could be identified in order to identify strategies for improving evaluator-client relationships and program outcomes.

In-Depth Investigations

The current study indicated several stress points in the evaluator-client relationship that merit further investigation such as the impact of changing funder requirements, grantee anxiety (particularly concerns about negative evaluation results), challenges to successful evaluation, threats to validity, the challenges inherent in active youth involvement in evaluation, and the unique role of internal evaluators. In-depth investigations of these stress points and strategies for addressing them would be of great interest to both evaluators and clients. For example, one area that merits further investigation is the changing role of the evaluator during different parts of the evaluation process or for different types of clients.

Questions resulting from implementation challenges to both program and evaluation methodologies describe a number of issues that would benefit from in-depth investigations. External forces impacting evaluations, fidelity problems with both program and evaluation methodologies, and difficulties recruiting and retaining youth program participants are three issues that come immediately to mind. The impact of current environmental factors on youth programs such as the increased competition for funds and increased demand for positive outcomes is another area ripe for investigation.

Implications for Programs Serving Youth with Disabilities

Programs that serve at-risk populations such as students with disabilities may be especially concerned with the results of this study. For example, the second National Longitudinal Transition Study (NLTS2) found significant differences in post-high school outcomes for students with disabilities and students without disabilities:

- “A considerable gap in achievement in language arts, mathematics, science, and social studies exists between youth with disabilities and their peers....From 77 percent to 86 percent of youth with disabilities [had] standard scores below the mean across subtests” (Wagner, Newman, Cameto, Levine, & Garza, 2006, p. 4).
- The high school completion rate for students with disabilities varied significantly by disability. For example, only 56% of youth with emotional disturbances reported completing high school, a much lower rate than youth without disabilities or those with other disabilities (Wagner et al, 2006, p. 7).
- The rate of enrollment in postsecondary school of youth with disabilities [was] about half that of their peers without disabilities (Wagner et al, 2006, p. 8).

- The rate of youth with disabilities who were employed was also “substantially below the 63 percent employment rate among same-age out-of-school youth in the general population” (Wagner et al, p. 8).

Given these data, programs serving youth with disabilities with the goals of improving high school outcomes are likely to have greater challenges meeting program goals than those serving youth without disabilities. Considering today’s climate of increased competition for funds and increased demands for positive outcomes this will likely create greater stress in the program which may carry over into the evaluator-client relationship. Evaluators may also experience longer learning curves in familiarizing themselves with the program’s context, disability etiquette, relevant disability laws and regulations, and other considerations that may add to their stress.

It is essential that programs serving youth with disabilities have valid evaluation results to document the effectiveness of their program interventions, even if the increments of change are smaller than programs for youth without disabilities. It is also essential that funders include needs assessments to map the extent of a problem, formative evaluations to improve or refine programs, and process evaluations to aid in implementation or replication of a program rather than relying solely on outcome or impact evaluations that focus on producing change or achieving specific outcomes. It is not implied that outcome or impact evaluations should not be used for programs serving youth with disabilities, but that the desired change, outcomes, and timeline should be appropriate and realistic. As noted previously, additional research studies on the evaluation of programs for youth with disabilities would also be beneficial.

Recommendations for Policy and Practice

Given that this study was exploratory and had the intent of identifying questions for future study, recommendations for policy, practice, and professional development will not be addressed in detail. However, it is evident from the study findings that some recommendations are in order.

Funders, particularly the federal government which was singled out by several survey respondents, should be made aware of the importance of the evaluator-client relationship in conducting valid and reliable evaluations that can be used to improve programs and program outcomes. Funders should ensure that their policies and procedures support the evaluator-client relationship and provide a context in which the relationship can succeed.

Some funders have excellent resources that support the evaluator-client relationship by providing common ground for grant implementation and communication between the two parties as noted in Appendix E. However, grant managers and other funder staff could do more to lay the ground work for effective evaluations by (a) requiring that evaluation planning, including evaluability determinations, be incorporated into program planning and the grant writing process, (b) providing adequate funding and resources for the evaluation piece, (c) encouraging evaluations that look beyond impact or summative evaluations to include lessons learned, best practices, program enhancements and improvements, and (d) recognizing that evaluations of programs serving at-risk youth populations may indicate that planned outcomes were not achieved-

-but still provide valuable information for improving services and outcomes for those groups in the future.

Funders could also support the exploration of the evaluator-client relationship, by funding and encouraging research based on the questions identified in this study on issues such as the psychology of evaluation, evaluator-client communications, interpersonal aspects of the relationship, and the day-to-day management of the evaluation process. Finally, funders could solicit the input of evaluators, youth service providers, youth, and other stakeholders when developing evaluation policy, guidelines, and grant solicitations. The American Evaluation Association's Evaluation Policy Task Force (EPTF) is making headway with the federal government in this regard by establishing connections with the Office of Management and Budget at the White House, providing advice on integrating evaluation into the Program Assessment Rating Tool, and discussing appropriate methods for impact evaluations. The Task Force is now expanding its contacts in the executive and legislative branches and has developed a series of handouts to support its work. (More information is available on the EPTF Web site at <http://www.eval.org/EPTF.asp>.)

Recommendations for Professional Development

Funders, youth service providers, and evaluators need to ensure that effective training and information related to the evaluator-client relationship is provided—if not required. It should be possible to build discussions of the relationship into postsecondary evaluation courses and professional development courses provided by evaluator associations. However professional development will be a challenge for youth service providers since funding is scarce and many staff members do not have a formal background in program management or implementation, in general, or evaluation, in

particular. Leaders of youth service organizations will need to provide time and support for such learning.

Youth service providers need to ensure that staff and appropriate stakeholders are competent and comfortable with the evaluation process and are prepared to develop and maintain effective working relationships with evaluators and funders to ensure proper planning and implementation of program evaluations. The National Collaborative on Workforce and Disability for Youth (NCWD/Youth) has developed KSAs (knowledge, skills, and abilities) needed for youth service professionals and is making headway in recognizing youth services as a profession. However, only a few references are made to evaluation skills or knowledge in the core KSA document (McCain, Gill, Wills & Larson, 2004):

- “Ability to evaluate and adjust programs based on outcome measurement and data” is included in Competency Area Program Design and Delivery (p. 9)
- Appendices A and B – Matrices of Competencies for Practitioners Serving Youth with Disabilities and for Practitioners Serving Youth with Disabilities describe 12 national certification programs of which only five specifically mention evaluation, program evaluation, or evaluation methods (pp. 26-27). None mention the evaluator/client relationship.

Colleges, universities, The Evaluators Institute, professional evaluation associations, and other groups that train and support evaluators need to do more to ensure that their policies, practices, and programs include discussions, resources, and professional guidelines that recognize and support the evaluator/client relationship.

Particular attention should be paid to preparing new evaluators to establish and nurture positive, mutually beneficial relationships with their clients.

An examination of the 2009 program schedule for The Evaluators Institute reveals only one course, “Working with Evaluation Stakeholders,” (offered at two of the three TEI sites) that is related to the evaluator/client relationship. However the course description, an excerpt of which follows, does make a strong case for the importance of the evaluator-client relationship—and makes a fitting ending for this discussion:

Working with stakeholders is a fact of life for evaluators. That interaction can be productive and beneficial to the conduct of a study that informs decisions that have positive outcomes for decision makers and program recipients. Or that interaction can be draining and often conflicting for both the evaluator and the stakeholders and lead to studies that are misguided, cost too much and take too long to do, and never get used; and to studies that never get done. So, this is an incredibly important topic for evaluators to explore. (The Evaluators Institute, n.d., Current Program Page).

References

- Abramson, A.J., Salamon, L.M. & Russell, J. (2006) *The nonprofit sector and the Federal Budget: Analysis of President Bush's FY 2007 Budget*. (Nonprofit Sector Research Fund Working Paper Series). Washington, DC: The Aspen Institute. Retrieved July 2, 2007, from http://www.nonprofitresearch.org/usr_doc/BushBudget.pdf
- ACT, Inc. (2007). *National collegiate retention and persistence to degree rates*. (Iowa City, IA: Author. Retrieved August 24, 2007, from http://www.act.org/path/policy/pdf/retain_2007.pdf
- Administration for Children & Families. (2003, December). *The program manager's guide to evaluation*. Washington, DC: U.S. Department of Health and Human Services. Retrieved April 8, 2006, from http://www.acf.hhs.gov/programs/opre/other_resrch/pm_guide_eval/index.html
- Alreck, P.L. & Settle, R.B. (1995). *The survey research handbook: Guidelines and strategies for conducting a survey*. (2nd ed.). New York, NY: McGraw-Hill.
- American Council on Education & General Educational Development Testing Service. (2007, November). *2006 GED testing program statistical report*. Washington, DC: Authors. Retrieved December 23, 2007, from <http://www.acenet.edu/bookstore/pdf/GEDASR06.pdf>
- American Evaluation Association. (2004, July). *Guiding principles for evaluators*, (Rev.) Fairhaven, MA: Author. Retrieved July 13, 2007, from <http://www.eval.org/Publications/GuidingPrinciplesPrintable.asp>
- American Evaluation Association. (n.d.). *About us*. Retrieved August 8, 2009, from

<http://www.eval.org/aboutus/organization/aboutus.asp>

American Youth Policy Forum. (1998, October 23). *Forum Brief: The educational and labor market performance of GED recipients*. Washington, DC: Author.

Retrieved December 23, 2007, from

<http://www.aypf.org/forumbriefs/1998/fb102398.htm>

Arenson, K.W., (2004, May 15). More students opt for G.E.D.—skirting high school hurdle. *The New York Times*. Retrieved December 23, 2007, from

http://performanceassessment.org/articles/pa_moreyouths.html

Ashby, C.M. (2007, April 18). *No Child Left Behind Act: Education actions may help improve implementation and evaluation of Supplemental Educational Services*.

(GAO -07-738T) Washington, DC: Government Accountability Office.

Ashby, C.M. (2008, February, 28). *Disconnected youth: Federal action could address some of the challenges faced by local programs that reconnect youth to education and employment*. (GAO-08-313). Washington, DC: Government Accountability Office.

Bain, O. (2003, February). CSR schools improve student achievement: Evidence from the meta-analysis. (Research Brief). Washington, DC: National Clearinghouse for Comprehensive School Reform.

Barkdoll, G.L. (1980, March-April). Type III evaluations: Consultations and consensus. *Public Administration Review*, 40, 174-179. [Quoted in Patton, 1997]

- Barton, N. & Wilhelm, I. (2007, April 5). Foundation assets grow sharply. *The Chronicle of Philanthropy*, 19, 12. Retrieved July 2, 2007, from <http://philanthropy.com/free/articles/v19/i12/12000701.htm>
- Barton, P. (2005, February). *One third of a nation: Rising dropout rates and declining opportunities*. (ETS Policy Information Report). Princeton, NJ: Educational Testing Service.
- Bell, S.H. (2001). Improving the evaluation of DOL/ETA pilot and demonstration projects: A guide for practitioners. (Research and Evaluation Report Series 01-A). Washington, DC: U.S. Department of Labor, Employment and Training Administration.
- Bellis, D.D. (2004, February 23). *Workforce Investment Act: Labor actions can help states improve quality of performance outcome data and delivery of youth services*. (GAO-04-308). Washington, DC: General Accounting Office.
- Bellis, D.D. (2005, October). Education's data management initiative: Significant progress made, but better planning needed to accomplish project goals. (GAO-06-06). Washington, DC: General Accounting Office.
- Bertoni, D. (2007, August 3). *Highlights of a forum: Modernizing federal disability policy*. (GAO-07-934SP). Washington, DC: Government Accountability Office.
- Blank, M.J., Melaville, A., & Shah, B. (2003, May). Making the difference: Research and practice in community schools. Washington, DC: Coalition for Community Schools, Institute for Educational Leadership.
- Boghossian, P. (2006). Behaviorism, constructivism, and Socratic pedagogy. *Educational Philosophy and Theory*, 38, 6, 713-722.

- Bond, S.L., Boyd, S.E., Rapp, K.A., Raphael, J., & Sizemore, B.A. (1997). *Taking stock: A practical guide to evaluating your own programs*. Chapel Hill, NC: Horizon Research, Inc.
- Branch-Smith, E., Gray, R., Fruchter, N., Hernandez, M., Joselowsky, F., Nichols-Solomon, R., Simmons, W., & Warren, C. (2006). *A framework for success for all students*. New York, NY: Carnegie Corporation of New York.
- Braskamp, L.A., Brandenburg, D.C., & Ory, J.C. (1987, Winter). Lessons about client expectations. *New Directions for Program Evaluation*, 36, 63-74.
- Brillinger, D.R. (2002, February). John Wilder Tukey (1915-2000). *Notices of the American Mathematical Society*, 49, 2, 193-201.
- Buechel, K.W., Keating, E.K., & Miller, C. (2007, March 15). *Capital ideas: Moving from short-term engagement to long-term sustainability*. Boston, MA & New York, NY: Hauser Center for Nonprofit Organizations, Harvard University & Nonprofit Finance Fund. Retrieved August 19, 2007, from <http://www.nonprofitfinancefund.org/docs/2007%20May%20onward/Capital%20Ideas/proceedings%20web.pdf>
- Carasso, A., Steuerle, C.E. & Reynolds, G. (2007). *Kids' share 2007: How children fare in the federal budget*. Washington, DC: Urban Institute. Retrieved July 2, 2007 from http://www.urban.org/UploadedPDF/411432_Kids_Share_2007.pdf
- Carasso, A., Steuerle, C.E., Reynolds, G., Vericker, T., & Macomber, J. *Kids' share 2008: How children fare in the federal budget*. Washington, DC: Urban Institute.
- Carey, K. (2004, May). *A matter of degrees: Improving graduation rates in four-year colleges and universities*. (A Report by the Education Trust). Washington, DC:

The Education Trust. Retrieved August 24, 2007, from
<http://www2.edtrust.org/NR/rdonlyres/11B4283F-104E-4511-BOCA-1D3023231157/0/highered.pdf>

Carman, J.G. (2007, March). Evaluation practice among community-based organizations: Research into the reality. *American Journal of Evaluation*, 28, 1, 60-75.

Carson Research Consulting. (2007, November). Basic principles of systems thinking. *The Catalyst*, 1-2.

Casner-Lotto, J. (2006, October). *Are they really ready to work?: Employers' perspectives on the basic knowledge and applied skills of new entrants to the 21st Century U.S. workforce*. New York, NY: The Conference Board, the Partnership for 21st Century Skills, Corporate Voices for Working Families, & the Society for Human Resource Management. Retrieved December 23, 2007 from
<http://www.conference-board.org/publications/describe.cfm?id=1218>

Cecil, D.P. (2004). *Diffusing the United Way model of outcome measurement from United Ways to human service programs*. (ProQuest Dissertations and Theses No. AAT 3142803). Columbia, SC: University of South Carolina.

Center for the Study of Social Policy. (1995, April). Building new futures for at-risk youth: Findings from a five year, multi-site evaluation. Washington, DC: Author. Retrieved June 6, 2006, from <http://www.cssp.org/publications/index.php>

Chambers, K. (2006, August 17). Nonprofits struggle finding needed funds. *JDNews.com*. Retrieved July 11, 2007, from
<http://www.jdnews.com/common/printer/view.php?db=jdn&id=4632>

- Charlton, J.I. (1998). *Nothing about us without us*. Berkeley, CA: University of California Press.
- Checkoway, B. & Richards-Schuster, K. (2003). Youth participation in community evaluation research. *American Journal of Evaluation*, 24, 1, 21-33.
- Chinman, M., Imm, P., & Wandersman, A. (2004). Getting to outcomes 2004: Promoting accountability through methods and tools for planning, implementation, and evaluation. (Technical Report TR-101-CDC). Santa Monica, CA: RAND Corporation. Retrieved April 8, 2006, from http://www.rand.org/pubs/technical_reports/TR101/
- Chudowsky, N., Chudowsky, V., & Kober, N. (2007, June). *Answering the question that matters most: Has student achievement increased since No Child Left Behind?* Washington, DC: Center on Education Policy.
- Community Research Associates, Development Associates, & Development Services Group, Inc. (2001) *Disproportionate minority confinement technical assistance manual*. (2nd ed., 2nd printing). Champaign, IL, Arlington, VA, & Bethesda, MD: Authors. Retrieved April 9, 2006 from http://www.ncjrs.gov/html/ojjdp/dmc_ta_manual
- Council on Foundations. (2009, February 2). Asset declines and investment strategy changes by family, independent, and public foundations. Arlington, VA: Author. Retrieved March 21, 2009, from <http://www.cof.org/files/Documents/Conferences/09FamPhilConf/EconSurvey4.pdf>

- Curran, B. (2005). Graduation counts: A report of the National Governors Association Task Force on state high school graduation data. Washington, DC: National Governors Association. Retrieved July 23, 2005, from <http://www.nga.org/portal/site/nga/menuitem.9123e83a1f6786440ddcbbeb501010a0/?vgnnextoid=7cadf891bc025010VgnVCM1000001a01010aRCRD>
- Dannels, S. (n.d.). *EDUC 330: Survey research methods*. Washington, DC: Graduate School of Education and Human Development, The George Washington University.
- Davis, D., Uhl, G., Barrington, T., Rowel, R., Squiers, L., Sharp, K., & O'Brien, G. (2004, October). Analysis of a database of materials for HIV prevention program evaluation. *Health Promotion Practice, 5*, 4, 444-450.
- Davis, G. (2005, November 20). *Sigma Xi postdoc survey methods*. (Sigma Xi Postdoc Survey Technical Report No. 2). Research Triangle Park, NC: Sigma Xi, The Scientific Research Society. Retrieved October 8, 1008 from http://postdoc.sigmaxi.org/news-items/results/tech_reports/survey-methods-tech-report-2.doc
- Donaldson, S.I., Gooler, L.E., & Scriven, M. (2002). Strategies for managing evaluation anxiety: Toward a psychology of program evaluation. *American Journal of Evaluation, 23*, 3, 261-273.
- Economic and Development Review Committee. (2007). *United States 2007*. (OECD Economic Surveys). Paris, France: Organisation for Economic Co-Operation and Development.

Education Portal. (2003-2007). A high school diploma V. the GED. [no location]:

Author. Retrieved December 23, 2007, from http://education-portal.com/articles/A_High_School_Diploma_v._the_GED.html

Ekstrand, L.E. (2001, October). *Juvenile justice: OJJDP reporting requirements for discretionary and formula grantees and concerns about evaluation studies.*

(GAO- 02-23). Washington, DC: General Accounting Office.

Ekstrand, L.E. & Rezmovic, E.L. (2003, September) *Justice outcome evaluations:*

Design and implementation of studies require more NIJ attention. (GAO-03-1091). Washington, DC: General Accounting Office.

Employment and Training Administration. (2007). Workforce system results: January

1- March 31, 2007. Washington, DC: U.S. Department of Labor. Retrieved 9/13/07 from

http://www.doleta.gov/performance/results/Quarterly_report/quarterly_report_03-31-07.pdf#page=13

Executive Office of the President of the United States, Office of Management and

Budget. (2007). *Budget of the United States, fiscal year 2008.* (Stock No. 041-

001-00635-4). Washington, DC: Government Printing Office. Retrieved July 3,

2007, from <http://bookstore.gpo.gov/collections/budget.jsp>

Fraenkel, J.R. & Wallen, N.E. (1996). *How to design and evaluate research in education.*

(3rd. ed.). New York, NY: McGraw-Hill, Inc.

Friedman, M. (2005). Trying hard is not good enough: How to produce measurable

improvements for customers and communities. Victoria, BC: Trafford

Publishing. Excerpts retrieved December 23, 2007, from

<http://www.amazon.com/Trying-Hard-Not-Good-Enough/dp/1412063973>

Gandal, M., Slover, L.M., Kraman, J., & Pache, T. (2004). The expectations gap: A 50-state review of high school graduation requirements. Washington, DC: Achieve, Inc.

Gilliam, A., Davis, D., Barrington, T., Lacson, R., Uhl, G., & Phoenix, U. (2002, June).

The value of engaging stakeholders in planning and implementing evaluations.

AIDS Education and Prevention, 14, Supplement A, 5-17.

Giving USA Foundation. (2009, June 10). U.S. charitable giving estimated to be \$307.65

billion in 2008. (News Release). Glenview, IL: Author. Retrieved June 10, 2009,

from http://www.givingusa.org/press_releases/gusa.cfm

Gladieux, L.E. & Swail, W.S. (1998). Financial aid is not enough: Improving the odds of college success. *The College Board Review*, 185, 1-11. Reprint retrieved

September 3, 2007, from

<http://www.educationalpolicy.org/pdf/Financial%20Aid%20is%20Not%20Enough.pdf>

Goodman Research Group. (2008, April). American Evaluation Association internal

scan report to the membership. Fairhaven, MA: American Evaluation

Association. Available at <http://www.eval.org/Scan/aea08.scan.report.pdf>

Gordon, M. (2009). Toward a pragmatic discourse of constructivism: Reflections on

lessons from practice. *Educational Studies*, 45, 39-58.

Greenstein, Robert. (2007, March 28). *Despite the rhetoric, budget would make nation's fiscal problems worse and further widen inequality*. Washington, DC: Center on

Budget and Policy Priorities. Retrieved July 3, 2007, from <http://www.cbpp.org/2-5-07bud.pdf>

- Hall, D. (2005, June). *Getting honest about grad rates: How states play the numbers game and students lose*. Washington, DC: The Education Trust.
- Hanson, S.G. (1998, September October). Evaluation: A tool for continuous improvement. *Knowledge Quest*, 27, 1, 44-45.
- Hasenfeld, Y. & English, R.A. (1974). *Human service organizations*. Michigan: University of Michigan Press. (p. 615) [Cited in Cecil, 2004]
- Hatry, H., van Houten, T., Plantz, M.C., & Taylor, M. (1996). *Measuring program outcomes: A practical approach*. (12th printing). Alexandria, VA: United Way of America.
- Havelock, R.G. & Hamilton, J.L. (2004). *Guiding change in special education: How to help schools with new ideas and practices*. Thousand Oaks, CA: Corwin Press
- Head, J., Sigritz, B., & Mazer, S. (2007, December). The fiscal survey of states. Washington, DC: National Governors Association and National Association of State Budget Officers. Retrieved December 11, 2007, from <http://www.nga.org/Files/pdf/FSS0712.PDF>
- Hicks, T. (2004, August 4). NCES seeks to standardize graduation rate data. *Education Daily*, 37, 145, 1.
- Hord, S.M., Rutherford, W.L., Huling-Austin, L., & Hall, G.E. (1989). *Taking charge of change*. Alexandria, VA: Association for Supervision and Curriculum Development.

- Horn, L. (2006, October). *Placing graduation rates in context: How 4-year college graduation rates vary with selectivity and the size of low-income enrollment*. (NCES-2007-161). Washington, DC: U.S. Department of Education, National Center for Education Statistics. Retrieved August 29, 2007, from <http://nces.ed.gov/pubs2007/2007161.pdf>
- Howell, E.M. & Yemane, A. (2006, June). An assessment of evaluation designs: Case studies of 12 large federal evaluations. *American Journal of Evaluation*, 27, 2, 219-236.
- Husch, B., Cheung, B., Cummings, L., Mazer, S., & Sigritz, B. (2009, June). *The fiscal survey of states*. Washington, DC: National Governors Association and National Association of State Budget Officers. Retrieved June 10, 2009, from <http://www.nasbo.org/Publications/PDFs/FSSpring2009.pdf>
- Hyslop-Margison, E.J. & Strobel, J. (2008, January-March). Constructivism and education: Misunderstandings and pedagogical implications. *The Teacher Educator*, 43, 1, 72-86.
- Innovation Center for Community and Youth Development & the Institute for Applied Research in Youth Development at Tufts University. (n.d.) *Planning and evaluation research center*. Located at <http://www.evaluationtools.org>
- Innovation Network, Inc. (2005a). *Evaluation plan workbook*. Washington, DC: Author. Retrieved August 16, 2007, from http://www.innonet.org/client_docs/File/evaluation_plan_workbook.pdf

Innovation Network, Inc. (2005b). *Logic model workbook*. Washington, DC: Author.

Retrieved August 16, 2007, from

http://www.innonet.org/client_docs/File/logic_model_workbook.pdf

Institute of Education Sciences. (2007, July 2). *IES awards \$62 million for longitudinal data systems*. (News Release). Washington, DC: Author.

Jackson, G. (2004, April 20). Seven key steps of program evaluation. *EDUC 281:*

Program evaluation. Retrieved May 10, 2008, from

<http://www.gwu.edu/~gjackson/281.html>

Jackson, G. (n.d.) Key considerations in evaluation design. *EDUC 281: Program*

evaluation. Washington, DC: Graduate School of Education and Human

Development, The George Washington University. Retrieved May 10, 2008, from

<http://www.gwu.edu/~gjackson/281.html>

Jarosewich, T., Essenmacher, V.L., Lynch, C.O., Williams, J.E., & Doino-Ingersoll, J.A.

(2006, Fall). Independent Consulting/Topical Interest Group: 2004 industry survey. *New Directions for Evaluation, 111*, 9-21.

Jaycox, L.H., McCaffrey, D.F., Ocampo, B.W., Shelley, G.A., Blake, S.M., Peterson,

D.J., Richmond, L.S., & Kub, J.E. (2006, September). Challenges in the

evaluation and implementation of school-based prevention and intervention

programs on sensitive topics. *American Journal of Evaluation, 27*, 3, 320-336.

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.

- Kegeles, S.M., Rebchook, G.M., & Tebbetts, S. (2005, August). Challenges and facilitators to building program evaluation capacity among community-based organizations. *AIDS Education and Prevention, 17*, 4, 284-299.
- W. K. Kellogg Foundation. (1998, January). *W.K. Kellogg Foundation evaluation handbook*. Battle Creek, MI: Author.
- W. K. Kellogg Foundation. (2001, December). *W.K. Kellogg Foundation logic model development guide*. Battle Creek, MI: Author.
- Kiernan, N.E., Kiernan, M., Oyler, M.A., & Gilles, C. (2005, June). Is a web survey as effective as a mail survey?: A field experiment among computer users. *American Journal of Evaluation, 26*, 2, 245-252
- Kingsbury, N. (2000, September). *Program evaluation: Studies helped agencies measure or explain program performance*. (GAO/GGD-00-204). Washington, DC: General Accounting Office.
- Kirsch, I., Braun, H., Yamamoto, K., & Sum, A. (2007, January). America's perfect storm: Three forces changing our nation's future. (Executive Summary/Policy Information Report. Princeton, NJ: Educational Testing Service. Retrieved July 13, 2007, from http://ets.org/Media/Education_Topics/pdf/ExecSummAmPerfectStorm.pdf
- Klein, A. (2009, May 20). Advocates fear for imperiled education programs. *Education Week, 28*, 33. Retrieved May 27, 2009 from <http://www.edweek.org>
- Kotter, J.P. (2006). Leading change: Why transformation efforts fail. (Harvard Business Review Reprint Number R0701J). Boston, MA: Harvard Business Review School Publishing Corporation. Retrieved March 19, 2007, from

<http://harvardbusinessonline.hbsp.harvard.edu/hbrsa/en/issue/0701/article/R0701JPrint.jhtml?path=arc>

Krueger, R. (2003, November 25). Response to RIN 1890-ZA00 Scientifically Based Evaluation Methods. Retrieved October 11, 2007, from

<http://www.eval.org/doecoverlet.htm> and <http://www.eval.org/doestatement.htm>

Kuss, L.J. (2002, June 26). Department of Labor, Office of Disability Employment Policy, Innovative demonstration grants for youth with disabilities. Federal Register, 67, 123, 43169-43181.

Laird, J., Kienzi, G., DeBell, M., & Chapman, C. (2007, June). *Dropout rates in the United States: 2005*. (NCES 2007-059). Washington, DC: U.S. Department of Education, National Center for Education Statistics. Retrieved July 6, 2007, from <http://nces.ed.gov/pubs2007/2007059.pdf>

Lawrence, S., Austin, A., & Mukai, R. (2007). *Foundation growth and giving estimates: Current Outlook*. (Foundation Today Series). New York, NY: Foundation Center. Retrieved July 2, 2007, from

<http://foundationcenter.org/gainknowledge/research/pdf/fgge07.pdf>

Lee, C.J., & Elizabeth, J. (2002, July 10). Foundations yank Pittsburgh school grants. *Pittsburgh Post-Gazette*. Retrieved April 9, 2006, from <http://www.post-gazette.com/printer.asp>

Levine, F.J. (2003, December 3). Response to RIN 1890-ZA00 Scientifically Based Evaluation Methods. Retrieved October 11, 2007, from

<http://www.eval.org/doesaera.htm>

- London, J.K., Zimmerman, K., & Erbstein, N. (2003). Youth-led research and evaluation: Tools for youth, organizational, and community development. *New Directions for Evaluation, 98*, 33-45.
- Loomis, C.J. (2006, June 25). Warren Buffet gives away his fortune. *Fortune Magazine*. Retrieved December 23, 2007, from <http://money.cnn.com/2006/06/25/magazines/fortune/charity1.fortune/index.htm>
- Lubienski, C. & Lubienski, S.T. (2006, August 30). Review of "On the public-private school achievement debate." (EPSL-0608-207-EPRU). Tempe, AZ: Education Policy Studies Laboratory, Arizona State University. Retrieved October 12, 2007, from <http://www.asu.edu/educ/epls/EPRU/ttreviews/EPSL-0608-207-EPRU.pdf>
- Manela, R.W. & Moxley, D.P. (1999). The new pillars of agency-based evaluation. *Applied Behavioral Science Review, 7*, 1, 41-57.
- Malloy, C.L. & Yee, P.A. (2006, Fall). Client relations: More than just "business". *New Directions for Evaluation, 111*, 67-77.
- Mathison, S. (2005). *Encyclopedia of evaluation*. [Excerpt]. Thousand Oaks, CA: Sage Publications, Inc.
- McCain, M., Gill, P., Wills, J., & Larson, M. (2004, May). *Knowledge, skills, and abilities of youth service practitioners: The centerpiece of a successful workforce development system*. Washington, DC: National Collaborative on Workforce and Disability for Youth, The Institute for Educational Leadership. Retrieved June 8, 2009, from <http://www.ncwd-youth.info/assets/background/ksa.pdf>
- McClintock, C. (2003). Commentary: The evaluator as scholar/practitioner/change agent. *American Journal of Evaluation, 24*, 1, 91-96.

- McMillan, J.H. & Schumacher, S. (1993). *Research in education: A conceptual introduction*. New York, NY: HarperCollins College Publishers.
- McNamara, C. (2007a). *Basic guide to outcomes-based evaluation for nonprofit organizations with very limited resources*. (Free Management Library) Minneapolis, MN & Toronto, Canada: Authenticity Consulting, LLC. Retrieved April 16, 2007, from <http://www.managementhelp.org/evaluatn/outcomes.htm>
- McNamara, C. (2007b). *Basic guide to program evaluation*. Minneapolis, MN & Toronto, Canada: Authenticity Consulting, LLC. Retrieved April 16, 2007, from http://www.managementhelp.org/evaluatn/fnl_eval.htm
- McNichol, E.C., Johnson, N., & Farkas, S. (2007, July 31). *State expenditure growth slowing*. Washington, DC: Center on Budget and Policy Priorities. Retrieved August 22, 2007, from <http://www.cbpp.org/7-31-07sfp.pdf>
- Mitra, A., Jain-Shukla, P, Robbins, A., Champion, H., & Durant, R. (2008). Differences in rate of response to web-based surveys among college students. *International Journal on E-Learning*, 7, 2, 265-281.
- Moody, R. (2003, December 3). Response to RIN 1890-ZA00 Scientifically Based Evaluation Methods. Retrieved October 11, 2007, from <http://www.eval.org/doe.nearesponse.pdf>
- Morabito, S.M. (2002, September). Evaluator Roles and Strategies for expanding evaluation process influence. *American Journal of Evaluation*, 23, 3, 321-330.
[Also quoted in Wiltz, 2005]

- Mulroy, E.A. & Lauber, H. (2004, October). A user-friendly approach to program evaluation and effective community interventions for families at risk of homelessness. *Social Work, 49*, 4, 573-586.
- Munley, M.E. (2003, December 3). Response to RIN 1890-ZA00 Scientifically Based Evaluation Methods. Retrieved October 11, 2007, from http://www.visitorstudies.org/pdfs/DOE_letter.pdf
- National Center for Education Statistics. (2007, July). Digest of education statistics: 2006. (NCES 2007-017). Washington, DC: U.S. Department of Education. Table 3 retrieved September 28, 2007, from http://nces.ed.gov/programs/digest/d06/tables/dt06_003.asp?referrer=list
- National Center for Public Policy and Higher Education. (2006). *Measuring up 2006: The national report card on higher education*. San Jose, CA: Author. (Retrieved August 24, 2007, from http://measuringup.highereducation.org/_docs/2006/NationalReport_2006.pdf
- National Collaborative on Workforce and Disability for Youth. (n.d.). *UDPATE: NCWD/Youth has released the Employer Guideposts for Success*. Washington, DC: Institute for Educational Leadership. Retrieved May 21, 2008, from http://www.ncwd-youth.info/resources_&_Publications/guideposts/
- National Collaborative on Workforce and Disability for Youth. (2003, May). *Evaluation guide for ODEP youth with disabilities grantees*. Washington, DC: Institute for Educational Leadership.

National Conference of State Legislatures. (2006, April 26). *State budget trends*. (NCSL Capitol Ideas). Retrieved August 22, 2007, from

<http://www.ncsl.org/programs/press/2006/cit060426.htm>

National Conference of State Legislatures. (2007, April 19). *State budgets post gains in FY 2007*. (Press Release). Retrieved August 22, 2007, from

<http://www.ncsl.org/programs/press/2007/pr070419budget.htm>

National Dropout Prevention Center for Students with Disabilities. (2006, April.) An analysis of state performance plan data for indicator 1 (Graduation). Clemson, SC: Author.

National Institute of Statistic Sciences & Education Statistics Services Institute

(NISS/ESSI). (2005). *National Institute of Statistic Sciences/Education Statistics*

Services Institute Task Force on Graduation, Completion, and Dropout Indicators

Final Report. (NCES 2005-105). Washington, DC: U.S. Department of

Education, National Center for Education Statistics. Retrieved February 17, 2005,

from <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2005105>

National Training Support Center. (n.d.) *ACRN evaluation workbook*. Washington, DC:

Author. Retrieved June 6, 2007, from

<http://www.acrnetwork.org/evaluation.htm#guides>

Neal, K.L. (2004). *Improving high school mathematics instruction: Using constructivist*

pedagogy. (ERIC No. ED490511). Fort Lauderdale, FL: Nova Southeastern

University.

New Commission on the Skills of the American Workforce. (2007). *Tough choices or*

tough times: The report of the New Commission on the Skills of the American

Workforce. (Executive Summary). Washington, DC: National Center on Education and the Economy

Nilsen, S.R. (2007, June 28). *Workforce Investment Act: Additional actions would further improve the workforce system*. (GAO-07-1051T). Washington, DC: Government Accountability Office.

Nonprofit Finance Fund. (2008, February 13). *NFF news: As the nonprofit sector faces recession, NFF outlines 5 recommendations to prepare for economic downturn*. New York, NY. Retrieved March 7, 2009, from <http://www.nonprofitfinancefund.org/news.php?autoID=113>

Office of Special Education Programs. (2007, June). U.S. Department of Education determination letters on state implementation of the IDEA. Washington, DC: U.S. Department of Education. Retrieved June 23, 2007, from www.ed.gov/fund/data/report/idea/partbspap/index.html

Olson, L. (2007). Linking learning to earning. In *From cradle to career: Connecting American education from birth to adulthood*. (Quality Counts 2007). Bethesda, MD: Editorial Policies in Education, Inc. Retrieved January 11, 2007, from <http://www.edweek.org/ew/articles/2007/01/04/17work-force.h26.html?print=1>

Orfield, G., Losen, D., Wald, J., & Swanson, C.B. (2004, March 11). *Losing our future: How minority youth are being left behind by the graduation rate crisis*. Cambridge, MA: The Civil Rights Project at Harvard University.

Organisation for Economic Co-operation and Development. (2006). *The OECD: Organisation for Economic Co-operation and Development*. Paris, France:

- Author. Retrieved December 22, 2007, from
<http://www.oecd.org/dataoecd/15/33/34011915.pdf>
- Orosz, J.J. (2001). Program evaluation in the foundation context. Grand Rapids, MI: Grand Valley State University, The Grantmaking School, Dorothy A. Johnson Center for Philanthropy and Nonprofit Leadership. Retrieved July 2, 2007, from
<http://philanthropy.com/free/articles//v19/i12/12000701.htm>
- Owen, J.M., Lambert, F.C., & Stringer, W.N. (1994). Acquiring knowledge of implementation and change: Essential for program evaluators? *Knowledge: Creation, Diffusion, Utilization* [now *Science Communication*] 15, 273-284.
- Paige, R. (2003, November 4). Scientifically based evaluation methods. *Federal Register*, 68, 213, 62445-62447.
- Paige, R. (2005, January 25). Scientifically based evaluation methods. *Federal Register*, 70, 15, 3586-3589.
- Partnership for 21st Century Skills. (n.d.) *21st Century skills outcomes*. Tucson, AZ: Author. Retrieved July 24, 2007, from
http://www.21stcenturyskills.org/index.php?option=com_content&task=view&id=254&Itemid=120
- Partnership for 21st Century Skills. (2004). *Mission statement*. Tucson, AZ: Author. Retrieved July 24, 2007, from
http://www.21stcenturyskills.org/index.php?option=com_content&task=view&id=188&Itemid=110
- Patton, M.Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Newberry Park, CA: SAGE Publications, Inc.

- Patton, M.Q. (1997). *Utilization-focused evaluation: The New Century text*. (3rd ed.). Thousand Oaks, CA: SAGE Publications, Inc.
- Patton, M.Q. (2008). *Utilization-focused evaluation*. (4th ed.). Thousand Oaks, CA: SAGE Publications, Inc.
- Peters, M. (2009, May 17). *The influence of classroom climate on student's mathematics self-efficacy and achievement: A multi-level analysis*. [Dissertation]. Washington, DC: The George Washington University Graduate School of Education and Human Development.
- Peterson, A.S. & Randall, L.M. (2006, August). Utilizing multilevel partnerships to build the capacity of community-based organizations to implement effective HIV prevention interventions in Michigan. *AIDS Education and Prevention*, 18, Supplement A, 83-95.
- Planty, M., Kena, G., & Hannes, G. (2009, June). The condition of education in 2009 in brief. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. Retrieved June 10, 2009, from <http://nces.ed.gov/pubs2009/2009082.pdf>
- Posner, P.L. & James, T. (2005, April). *Grants management: Additional Actions needed to streamline and simplify processes*. (GAO 05-335). Washington, DC: Government Accountability Office.
- Public Law 103-62, Government Performance and Results Act of 1993*, 107 Stat. 285
- Public Law 106-107, Federal Financial Assistance Management Improvement Act of 1999*, 113 Stat. 1488
- Public Law 107-110, No Child Left Behind Act of 2001*, 115, Stat. 1425.

- Public Law 107-279, Education Sciences Reform Act of 2002*, 116 Stat. 1940.
- Rallis, S.F., & Rossman, G.B. (2000). Dialogue for learning: Evaluator as critical friend. *New Directions for Evaluation*, 86, 81-92. [Quoted in Wiltz, 2005]
- Renger, R. (2006, December). Consequences to federal programs when the logic-modeling process is not followed with fidelity. *American Journal of Evaluation*, 27, 4, 452-464.
- Ritter, L.A. & Sue, V.M. (2007, Fall). Using online surveys in evaluation. *New Directions for Evaluation*, 115. [Note: the entire issue was written by the two authors.]
- Robelen, E.W. (2009, March 18). Education philanthropy catching a chill as economy cools charitable giving. *EducationWeek*, 28, 25, 1, 10-11.
- Robertson, R.E. (2005, September 23). *Vocational Rehabilitation: Better measures and monitoring could improve the performance of the VR program*. (GAO-05-865). Washington, DC: General Accounting Office.
- Robertson, R.E. (2006, June 20). *Summary of a GAO conference: Helping California youths with disabilities transition to work or postsecondary education*. (GAO-06-759SP). Washington, DC: Government Accountability Office.
- Robertson, R.E. (2007, January). *Federal disability assistance: Stronger federal oversight could help assure multiple programs' accountability*. (GAO- 07-236) Washington, DC: Government Accountability Office.
- Rogers, G. (n.d.). *Sample protocol for pilot testing survey items*. Baltimore, MD: ABET, Inc. Retrieved October 8, 2008, from

http://www.abet.org/Linked%20Documents-UPDATE/Assessment/Protocol_item%20testing.pdf

- Rossi, P.H., Lipsey, M.W., & Freeman, H.E. (2004). *Evaluation: A systematic approach*. (7th ed.). Thousand Oaks, CA: SAGE Publications.
- Rothstein, L.F. (1990). *Special education law*. (2nd ed.). White Plains, NY: Longman Publishers USA.
- Sabo, K. (2003, Summer). Editor's Notes [Special Issue on Youth Participatory Evaluation: A Field in the Making]. *New Directions for Evaluation*, 98, 1-11.
- Samuels, C.A. (2009, May 20). Restrictions on Spec.Ed. spending prompts complaints. *Education Week*, 28, 32, 7. Retrieved May 27, 2009 from <http://www.edweek.org>
- Sanders, J.R. (2003, Fall). Mainstreaming evaluation. *New Directions for Evaluation*, 99, 3-6)
- Saunders, M. (2000). Beginning an evaluation with RUFDATA: Theorizing a practical approach to evaluation planning. *Evaluation*, 6, 1. 7-21.
- Scheuren, F. (2004, June). What is a survey. Chicago, IL: National Opinion Research Center, University of Chicago.
- Schoenfeld, A.H. (2006a, March). What doesn't work: The challenge and failure of the What Works Clearinghouse to conduct meaningful reviews of mathematics curricula. *Educational Researcher*, 35, 2, 13-21. Retrieved October 8, 2007, from http://www.aera.net/uploadedFiles/Publications/Journals/Educational_Researcher/3502/03ERv35n2_Schoenfeld.pdf
- Schoenfeld, A.H. (2006b, March). Reply to comments from the What Works Clearinghouse on "What doesn't work". *Educational Researcher*, 35, 2, 23.

- Schwandt, T.A. (2005, March). The centrality of practice to evaluation. *American Journal of Evaluation*, 26, 1, 95-105.
- Schwinn, E. (2007, July 2). Federal budget situation for charities improves, study finds. *The Chronicle of Philanthropy News Update*. Retrieved July 3, 2007, from <http://philanthropy.com/news/updates/index.php?id=2623>
- Scott, G.A. (2009, February 26). *Workforce Investment Act: Labor has made progress in addressing areas of concern, but more focus needed on understanding what works and what doesn't*. (GAO-09-396T). Washington, DC: Government Accountability Office.
- Scriven, M. (1991). Beyond formative and summative evaluation. In M.W. McLaughlin and D.C. Phillips (eds.), *Evaluation and education: At quarter century*. Chicago, IL: University of Chicago Press. [Quoted in Wiltz, 2005]
- Scriven, M. (1996). Types of evaluation and types of evaluator. *Evaluation Practice*, 17, 2, 151-161.
- Shaul, M.S. (2005, September). *No Child Left Behind: Education could do more to better define graduation rates and improve knowledge about intervention strategies*. (GAO-05-879). Washington, DC: Government Accountability Office.
- Shaul, M.S. (2006, February). *Discretionary grants: Further tightening of Education's procedures for making awards could improve transparency and accountability*. (GAO-06-268). Washington, DC: Government Accountability Office.
- Shaul, M.S., Edwards, D., and Fucile, T. (2004, September). *Special Education: Improved timeliness and better use of enforcement actions could strengthen*

- education's monitoring system.* (GAO-04-879). Washington, DC: Government Accountability Office.
- Shumer, R. (2007). *Youth-led evaluation: A guidebook.* Clemson, SC: National Dropout Prevention Center, College of Health, Education, and Human Development, Clemson University.
- Simon, S. (2008, July 8). Design and analysis of pilot studies. Kansas City, MO: The Children's Mercy Hospital. Retrieved October 8, 2008, from <http://www.childrens-mercy.org/stats/weblog2004/PilotStudy.asp>
- Stoecker, R. (2007, December). The research practices and needs of non-profit organizations in an urban center. *Journal of Sociology & Social Welfare*, 34, 4, 97-119.
- Stufflebeam, D. (1994). Empowerment evaluation, objectivist evaluation, and evaluation standards: Where the future of evaluation should not go and where it needs to go. *Evaluation Practice*, 15, 3, 321-338. [Quoted in Wiltz, 2005]
- Sum, A., Harrington, P., Bartishevich, C., Fogg, N., Khatiwada, I., Motroni, J., Palma, S., Pond, N., Tobar, P., & Trub'skyy, M. (2003, February). *The hidden crisis in the high school dropout problems of young adults in the U.S.: Recent trends in overall school dropout rates and gender differences in dropout behavior.* Boston, MA: Center for Labor Market Studies, Northeastern University.
- SurveyMonkey. (2008). *Smart survey design.* Menlo Park, CA, & Portland, OR: Author.
- Swanson, C.B. (2007, June 12). Learning and earning. In *Diplomas count: Ready for what? Preparing students for college, careers and life after high school.*

- Bethesda, MD: Editorial Policies in Education, Inc. Retrieved June 14, 2007, from <http://www.edweek.org/ew/articles/2007/06/12/40jobs.h26.html?print=1>
- Taut, S.M. & Alkin, M.C. (2003). Program staff perceptions of barriers to evaluation implementation. *American Journal of Evaluation*, 24, 2, 213-226.
- Thayer, C.E. (2006) *Participatory evaluation in nonprofit organizations: Rhetoric or reality?* (ProQuest Dissertations and Theses No. AAT 3215660). Washington, DC: The George Washington University.
- The Barefoot Collective. (2009). The barefoot guide to working with organisations and social change. Capetown, SA: Author. Retrieved June 2, 2009, from <http://www.barefootguide.org/download.php>
- The Evaluators Institute, (n.d.) *2009 class schedule*. Washington, DC: The George Washington University. Retrieved June 8, 2009, from http://tei.gwu.edu/current_program.cfm
- The Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards* (2nd ed.) Thousand Oaks, CA: Sage.) Summary of the Standards retrieved September 14, 2007, from <http://www/eval.org/EvaluationDocuments/progeval.html>
- Tilson, G.P. Jr. (1986). *The status of evaluation within federally supported postsecondary model demonstration projects serving the school-to-work transition needs of disabled youth and young adults: An exploratory study*. (ProQuest Dissertations and Theses No. AAT 8627901). Washington, DC: The George Washington University.

- Timmons, J., Podmostko, M., Bremer, C., Lavin, D., & Wills, J. (2005). *Career planning begins with assessment: A guide for professionals serving youth with educational & career development challenges (Rev. Ed.)*. Washington, D.C.: National Collaborative on Workforce and Disability for Youth, Institute for Educational Leadership. Available at http://www.ncwd-youth.info/resources_&_Publications/manuals.php
- Toppo, G. (2007, April 10). Usefulness of education research questioned. *USA Today*. Retrieved October 8, 2007, from http://www.usatoday.com/news/education/2007-04-10-education-science_N.htm
- Tukey, J.W. (1980, February). We need both exploratory and confirmatory. *American Statistical Association*, 34, 1, 23-25.
- Tyler, J.H. (2004). The General Educational Development (GED) credential: History, current research, and directions for policy and practice. *Review of Adult Learning and Literacy*, 5, 3, 45-84. Retrieved December 23, 2007, from http://www.ncsall.net/fileadmin/resources/ann_rev/rall_v5_ch3.pdf
- U.S. Department of Education. (2009, April 13). Modifications to the questions in the April 2009 guidance on the Individuals with Disabilities Education Act, Part B. Washington, DC: Author. Retrieved May 27, 2009, from <http://www.ed.gov/policy/gen/leg/recovery/guidance/idea-guidance-mod.pdf>
- U.S. Department of Education. (2009, May 7). FY2010 budget summary: Section IV. Programs proposed for elimination. Washington, DC: Author. Retrieved May 11, 2009 from <http://www.ed.gov/about/overview/budget/budget10/summary/edlite-section4.html>

U.S. Office of Management and Budget. (n.d.a). ExpectMore.gov. Retrieved May 13, 2007, from <http://www.whitehouse.gov/omb/expectmore>

U.S. Office of Management and Budget. (n.d.b). U.S. Department of Education 2010 Budget Fact Sheet. Retrieved June 10, 2009, from http://www.whitehouse.gov/omb/fy2010_department_education

U.S. Office of Management and Budget. (n.d.c). U.S. Department of Labor 2010 Budget Fact Sheet. Retrieved June 10, 2009, from http://www.whitehouse.gov/omb/fy2010_department_labor

United Way of America. (2005, March). *Impact strategies, resources and results*. (United Way standards of excellence, Key area of operation 2: Aspire). Fairfax, VA: Author. Retrieved July 6, 2007, from <http://national.unitedway.org/about/soe.cfm>

Viadero, D. (2006, September 26). 'One Stop' research shop seen as too slow to yield views that educators can use. *Education Week*, 26, 5, 8-9. Retrieved October 18, 2007, from <http://www.edweek.org/ew/articles/2006/09/27/05whatworks.h26.html?print=1>

Viadero, D., (2009, April 1). No effects' studies raising eyebrows. *Education Week*, 28,27,1,14-25. Retrieved May 3, 2009, from http://www.edweek.org/ew/articles/2009/04/01/27rct_ep.h28.html

Wagner, M., Newman, L., Cameto, R., Levine, P, & Garza, N. (2006). *An overview of findings from Wave 2 of the National Longitudinal Transition Study-2 (NLTS2)*. (NCSE 2006-3004). Menlo Park, CA: SRI International. Retrieved December 20, 2007, from http://www.nlts2.org/reports/2006_08/index.html

- Wainer, H. (1978, Winter). Book reviews: Exploratory Data Analysis. *Applied Psychological Measurement*, 2, 1, 151-155.
- Walker, D. & Lambert, L. (1995). Learning and leading theory: A century in the making. In L. Lambert, D. Walker, D.P. Zimmerman, J.E. Cooper, M.D. Lambert, M.E. Gardner, & P.J.F. Slack,(Eds), *The constructivist leader* (pp. 1-27). New York, NY: Teachers College Press.
- Wandersman, A. & Snell-Johns, J. (2005, September). Empowerment evaluation: Clarity, dialogue, and growth. *American Journal of Evaluation*, 26,3, 421-428. Retrieved June 16, 2009, from <http://aje.sagepub.com>
- Weidenthal, C. (2002). An investigation of the implementation of the transition requirements of the Individuals with Disabilities Education Act of 1997 for early adolescents with learning disabilities. (Publication No. AAT3045495). (ProQuest Document ID No. 726397031). Washington, DC: The George Washington University.
- Weiss, C.H. (1998a). *Evaluation* (2nd ed.). Upper Saddle River, NJ: Prentice-Hall, Inc.
- Weiss, C.H. (1998b). Have we learned anything new about the uses of evaluation? *American Journal of Evaluation*, 19, 1, 21-33
- Welner, K.G. & Molnar, A. (2007, February 28). Truthiness in education. (Commentary). *Education Week*, 26, 25, 32 & 44. Retrieved March 1, 2007, from <http://www.edweek.org>.
- What Works Clearinghouse. (2007). *Dropout prevention topic report*. Washington, DC: Institute of Education Sciences, U.S. Department of Education. Retrieved August 20, 2007, from <http://www.whatworks.ed.gov/TopicReportLinks.asp?tid=6>

- Wheatley, M. (2002). When change is out of our control. In M. Effron, R. Gandossy, & M. Goldsmith (Eds.), *Human resources in the 21st century*. Hoboken, NJ: John Wiley & Sons, Inc. Retrieved May 31, 2009, from <http://www.margaretwheatley.com/articles/whenthechangeisoutofcontrol.html>
- Wheatley, M. (2006). *Relationships: The basic building blocks of life*. Retrieved May 31, 2009, from <http://www.margaretwheatley.com/articles/relationships/html>
- Wiltz, L.K. (2005, Winter). I need a bigger suitcase: The evaluator role in nonformal education. *New Directions for Evaluation*, 108, 13-28.
- Wons, E. (1999). Organizational change: An ethical, means-based approach. Minnetonka, MN: JPC Training & Consulting, LLC. Retrieved December 24, 2007, from <http://www.jpc-training.com/change/review.htm>

Appendix A

Survey Cover E-mail and Survey Instrument

To: [Individual E-mail Address]
 Date: December 3 and 18, 2008
 From: [Researcher E-mail Address]
 Subject: *Reminder: Survey on the Program Evaluator/Client Relationship*

[Italics indicate language that was added to the December 18 reminder and survey.]

This is a reminder that you are invited to participate in an IRB-approved research study on the relationship between program evaluators and their clients who serve youth within the age range of 14-25 years. Please respond on or before January 1, 2009.

Survey data will expand understanding of this critical relationship and will be shared with the field to expand the knowledge base. The study grew out my experiences as part of a team providing evaluation technical assistance to over 20 grantees who worked with youth transitioning to adulthood. The grantees ranged from very small local programs to statewide systems change coalitions and reflected a wide range of knowledge, experience, and comfort with evaluation—and evaluators.

The on-line survey will take approximately 15 minutes to complete. If you prefer a paper survey, please respond to this e-mail with your complete street address. A paper survey will be sent to you with a self-addressed, stamped return envelope. To ensure confidentiality, your e-mail or name and address will be kept separate from the data.

To be eligible for this study, you must:

- * provide evaluation services in the United States, and
- * have been an evaluator, evaluation consultant, or evaluation team member (including internal staff)
 - in the last five years
 - to an organization serving youth within the age range of 14-25 years
 - funded by government or foundation grant(s) for which the youth organization voluntarily applied.

Your responses are greatly appreciated. If you meet all the criteria and wish to participate, please access the survey through this link: [Secure Link]

This link is uniquely tied to this survey and your email address. Please do not forward this message.

If you do not wish to participate or receive further e-mails from me, please click the link below, and you will be automatically removed from the mailing list: [Secure Link]

Thank you,

Mary Podmostko
AEA Member
Doctoral Student
The George Washington University

You are receiving this email as a member of AEA. If you have concerns about the survey and would like to express them to the AEA leadership, please email them to aea@eval.org. Any concerns raised will be shared, confidentially, with the Executive Committee of the association. AEA allows its membership list to be used up to twice a year for research, reviewed by the Executive Committee, that focuses on the field of evaluation. If you would like to opt-out of AEA's research list, please send an email request to heidi@eval.org. Please note that we encourage you to consider remaining on the list as such research strengthens and furthers the field's knowledge base.

AEA Evaluators and Youth Programs

Welcome to the Survey on the Program Evaluator/Client Relationship

Filling out the survey is entirely voluntary, and you may stop your participation in this study at any time. Your answers will be treated as confidential and will be reported only in the aggregate. No serious risks to respondents are anticipated. You will not benefit directly from your participation in the study, but the survey results will be submitted to the American Journal of Evaluation for publication to inform the profession. You may request a copy of the final report by contacting me at the e-mail or telephone number below.

Your cooperation is greatly appreciated. Your willingness to participate in this doctoral study (*IRB #050857*) is implied if you proceed with completing the survey. If you have any additional questions, please feel free to call GWU's Office of Human Research at 202-994-2715. You may also contact me at 904-225-9718 or maryp@gwu.edu.

Please base your answers on your MOST RECENT role as an

- evaluator, evaluation consultant, or evaluation team member (including internal staff)
- for a grant-funded program in the United States
- serving youth within the age range of 14-25 years.

Please respond for grant-funded programs where the youth program voluntarily applied for a grant, thereby indicating agreement with evaluation requirements. Do not respond for formula-funded programs where funds were received from predetermined allocations such as those in the No Child Left Behind or Workforce Investment Acts.

Thank you!

Mary Podmostko
AEA Member
Doctoral Student
The George Washington University

AEA Evaluators and Youth Programs

Section I. Youth Program Information

Please provide descriptive Information on the MOST RECENT grant-funded youth program for which you provided evaluation services in the last five years

You will have an opportunity at the end of the survey to provide additional information regarding your answers if you so desire.

1. What best describes the TYPE of youth organization you served? [Forced choice]

Please use the "Other" category if (a) your answer is not found in the list provided or (b) you cannot identify just one response and wish to select two or more answers.

- a. Individual nonprofit/community-based organization or 501(c)(3)
- b. Public school or school district or component thereof
- c. Four year college or university
- d. Community or junior college
- e. Government agency
- f. Collaboration of some or all of the above groups
- g. Other (please specify) _____

2. What was the PRIMARY FUNDING SOURCE for the grant program that was evaluated? [Forced choice]

Please use the "Other" category if (a) your answer is not found in the list provided or (b) you cannot identify just one response and wish to select two or more answers.

- a. Federal government
- b. State government
- c. Local/municipal government
- d. Foundation
- e. Corporation
- f. Other (please specify) _____

3. What type of youth (within the age range 14-25) was the PRIMARY TARGET of the grant program? [Forced choice]

- a. Academically at-risk youth
- b. Academically gifted youth
- c. Low-income youth
- d. Minority youth
- e. English Language Learners
- f. Youth with disabilities
- g. Gay/lesbian/bisexual/transgender youth

h. All youth within the geographic scope of the grant

Other (Please specify) _____

4. What type of youth outcome was the PRIMARY FOCUS of the grant program?
[Yes, No, Not applicable/Don't know]

- a. High school completion/graduation
- b. Career development/employment
- c. College enrollment/completion
- d. Independent living
- e. Healthy choices

Other (please specify) _____

5. What was the PRIMARY REASON that the youth organization hired an evaluator or consultant?
[Force choice]

Please use the "Other" category if (a) your answer is not found in the list provided or (b) you cannot identify just one response and wish to select two or more answers.

- a. Lack of organizational capacity to conduct an evaluation
- b. Funder requirement
- c. External requirement such as accreditation
- e. Stakeholder decision
- f. Board of directors/senior management decision
- g. Other (please specify) _____

AEA Evaluators and Youth Programs

Section II. Your Role

Please describe your role as evaluator, evaluation consultant, or evaluation team member for the most recent grant-funded youth program for which you provided services.

Be sure to respond to each item in each Yes/No and Likert question. If you select "other", you must also fill in the blank.

You will have an opportunity at the end of the survey to provide additional information regarding your answers if you so desire.

6. For what type of organization did you work when providing evaluation services to the youth organization? [Forced choice]

- a. The youth organization being evaluated
- b. Yourself (i.e., self-employed consultant)
- c. Independent consulting firm
- d. Nonprofit services organization/agency
- e. College or university
- f. Government agency
- g. Corporate or private foundation
- h. Other (please specify) _____

7. What BEST DESCRIBES your role in providing evaluation services to the youth organization? [Yes, No, Not applicable/Don't know]

- a. Judge of program merit or worth
- b. Auditor, inspector, or compliance/accountability evaluator
- c. Researcher or scientist
- d. Evaluation facilitator or guide
- e. Team member with evaluation perspective or internal advisor
- f. Collaborator or partner
- g. Program improvement consultant
- h. Empowerment facilitator
- i. Advocate for social change or supporter of a cause
- j. Educator or capacity builder
- k. Organization development consultant or change agent
- l. Program advocate
- m. Critical friend

Other (please specify) _____

8. Did you encounter staff and/or stakeholder ANXIETY related to their:

[Likert scale: None at all, A small amount, A moderate amount, A large amount, Not applicable/Don't know]

- a. Lack of experience/knowledge with evaluation
- b. Negative past experiences with evaluation
- c. Fear of failure, negative results, or consequences
- d. Organizational culture or social norms resistant to evaluation
- e. Role conflicts between evaluator(s) and stakeholders
- f. Concerns about evaluator competence (technical or social)
- g. Concerns about evaluator understanding of program environment

Other (please specify) _____

Section III. Planning and Implementing of the Above Referenced Evaluation

Please describe the planning and implementation of the most recent grant-funded program serving youth within the age range 14-25 years for which you provided services.

You will have an opportunity at the end of the survey to provide additional information regarding your answers if you so desire.

9. What was the PRIMARY FOCUS of the evaluation? [Forced choice]

Please use the "Other" category if (a) your answer is not found in the list provided or (b) you cannot identify just one response and wish to select two or more answers.

- a. Needs assessment to verify and map the extent of a problem
- b. Formative evaluation to improve and refine the program
- c. Process evaluation to aid program implementation and replication
- d. Outcome/impact evaluation to assess program effectiveness in producing change
- e. Summative evaluation to assess program quality/impact for accountability and decision-making purposes
- f. Other (please specify) _____

10. Did the evaluation planning process include determinations of [Yes, No, Not applicable/Don't know]

- a. Program evaluability (i.e., determination of program readiness for evaluation)
- b. Contract of agreement terms
- c. Evaluation focus
- d. Methodology (research design)
- e. Evaluation field test (of materials, instruments, processes, etc.)
- f. Data collection
- g. Data editing/reporting
- h. Reporting

Other (please specify) _____

11. Were youth actively involved in the evaluation planning and implementation process (i.e., in roles other than program participant)? [Likert scale]

- a. Youth were actively involved

12. What challenges to the implementation of the evaluation plan perhaps had to be overcome? [Likert scale]

- a. Rigid organizational procedures or settings
- b. Inadequate/Problematic data
- c. Disagreement/Debate over who should conduct the evaluation

- d. Lack of staff cooperation, collaboration, or communication
- e. Limited methodological capacity of staff
- f. Lack of staff resources such as time or funding
- g. Unexpected/Unanticipated evaluation results or consequences
- h. Friction between youth organization and funder
- i. Misapplication of theories (e.g. constructionism or post-modernism)

Other (please specify) _____

13. What threats to the validity of the evaluation plan were encountered?

[Likert scale]

- a. Outside forces affected intervention or results
- b. Interventions mutated from the original design
- c. Youth reacted to please or displeasure program operators (halo/cloud effects)
- d. Youth became sensitized to intervention measures over time
- e. Maturation of youth participants biased the observed results
- f. Evaluation process caused disruptions that impacted results
- g. Measurements became unstable or changed over time
- h. Control groups were not comparable (initially or over time)
- i. Program interventions migrated to the control group

Other (please specify) _____

14. What evidence of successful evaluation implementation and usefulness were perhaps observed?

[Likert scale]

- a. Improved stakeholder focus, ownership, and/or skills
- b. Improved organizational capacity, impact, viability, and/or credibility
- c. Improved program implementation
- d. Improved program results
- e. Improved use of evaluation results by youth program, funder, or other audience

Other (please specify) _____

15. With whom were evaluation results shared? [Yes, No, Not applicable/Don't know]

- a. Program sponsors and/or funders
- b. Program director and staff
- c. Program participants
- d. Practitioners in the youth service field
- e. Policymakers and/or opinion leaders
- f. Social scientists
- g. Other evaluators
- h. The general public

Other (please specify) _____

Section IV. Final Section

Final questions about the most recent grant-funded youth program serving youth within the age range 14-25 years for which you provided services.

16. Please describe a serious implementation challenge or threat to validity that you encountered and successfully dealt with during the afore-mentioned evaluation. What was the challenge or threat, how did you deal with it, and with what effect?

17. What evaluation resources (guides, publications, Web sites, etc.) do you find to be particularly helpful or effective in working with youth organizations (not just the youth program described previously)?

18. Use this question to make any comments of your choosing. For example, was there something notable about the evaluator/client relationship that was not capture in the previous survey questions? Was the client described representative of other youth organizations you have served? Would you like to explain or expand upon any of your previous answers?

The End

You have completed the survey. Thank you!

Appendix B

Supporting Literature for Survey Questions

Section I. Descriptive Information on the *MOST RECENT* grant-funded youth program for which you provided evaluation services in the last five years

1. What best describes the *type* of youth organization you served?
 - h. Individual nonprofit/community-based organization/501(c)(3)
 - i. Public school or school district or component thereof
 - j. Four year college/university
 - k. Community/junior college
 - l. Government agency
 - m. Collaboration of some or all of the above groups
 - n. Other _____

2. What was the *primary funding source* for the grant program that was evaluated?
 - g. Federal government
 - h. State government
 - i. Local/municipal government
 - j. Foundation
 - k. Corporation
 - l. Other _____

3. What type of youth (ages 14-25) was the *primary target* of the grant program?
 - i. Academically at-risk youth
 - j. Academically gifted youth
 - k. Low-income youth
 - l. Minority youth
 - m. English Language Learners
 - n. Youth with disabilities
 - o. Gay/Lesbian/Bisexual/Transgender youth
 - p. All youth within the geographic scope of the grant
 - q. Other _____

4. What type of youth outcome was the *primary focus* of the grant program?
 - f. High school completion/graduation
 - g. Career development/employment
 - h. College enrollment/completion
 - i. Independent living
 - j. Healthy choices
 - k. Other _____

5. What was the *primary reason* the youth organization hired an evaluator or consultant?
 - a. Lack of organizational capacity (Thayer, 2006)
 - b. External requirements such as accreditation (Thayer, 2006)
 - c. Funder requirement (Kegeles, Rebchook, & Tebbetts, 2005; Rossi, Lipsey & Freeman, 2004; Manela & Moxley, 1999)
 - d. Stakeholder decision
 - e. Board of directors/senior management decision (Thayer, 2006)
 - f. Add credibility to the evaluation
 - f. Other _____

Section II. Your Role as Evaluator or Evaluation Consultant on the most recent grant-funded youth program for which you provided services.

6. For what type of organization did you work when providing evaluation services to the youth organization?
 - i. Corporate or private foundation
 - j. Government agency
 - k. College or university
 - l. Independent consulting firm
 - m. Yourself (Self-employed consultant)
 - n. The youth organization being evaluated
 - o. Other: _____
7. What *best describes* your role with the youth organization?
 - a. Judge of program merit or worth (Scriven, 1991; Stufflebeam, 1994; Patton, 1997)
 - b. Auditor, inspector, compliance/accountability evaluator (Barkdoll, 1980; Patton, 1997)
 - c. Researcher/scientist (Barkdoll, 1980; Weiss, 1998a; Patton, 1997)
 - d. Evaluation facilitator/guide (Weiss, 1998a; Patton, 1997; Wiltz, 2005)
 - e. Team member with evaluation perspective, internal advisor (Patton, 1997, Wiltz, 2005)
 - f. Collaborator/partner (Barkdoll, 1980; Weiss, 1998a; Patton, 1997; Wiltz, 2005)
 - g. Program improvement consultant (Patton, 1997)
 - h. Empowerment facilitator (Weiss, 1998a; Patton, 1997)
 - i. Supporter of cause, advocate for social change (Patton, 1997; Morabito, 2002)
 - j. Educator, capacity builder (Morabito, 2002. Wiltz, 2005; Weiss, 1998a)
 - k. Organization development consultant, change agent (Owen, Lambert & Stringer, 1994; McClintock, 2003; Wiltz, 2005)
 - l. Program advocate (Wiltz, 2005)
 - m. Critical friend (Rallis & Rossman, 2000; Weiss, 1998a)
 - n. Other _____

8. Did youth organization staff or stakeholders exhibit anxiety related to
 - a. lack of experience with or knowledge of with evaluation (Donaldson, Gooler, & Scriven, 2002; Taut & Alkin, 2003; Kegeles, Rebchook, & Tebbetts, 2005)
 - b. negative past experiences with evaluation ((Kegeles, Rebchook, & Tebbetts, 2005; Donaldson, Gooler, & Scriven, 2002; Taut & Alkin, 2003; Shadish, Cook & Leviton, 1991)
 - c. fear of negative evaluation results or consequences (Donaldson, Gooler, & Scriven, 2002: Taut & Alkin, 2003; Kegeles, Rebchook, & Tebbetts, 2005)
 - d. organizational culture or social norms resistant to evaluation (Donaldson, Gooler, & Scriven, 2002; Weiss, 1998a)
 - e. role conflicts between evaluator and stakeholders (Donaldson, Gooler, & Scriven, 2002)
 - f. concerns about evaluator's competence (technical or social) (Braskamp, Brandenburg, & Ory, 1987; Taut & Alkin, 2003)
 - g. distrust of the evaluator (Gilliam, Davis, Barrington, Lacson, Uhl, & Phoenix, 2002)
 - h. concerns about evaluator's understanding of the program environment (Braskamp, Brandenburg, & Ory, 1987; Taut & Alkin*, 2003)
 - i. other _____

Section III. Planning and Implementation of the Above Referenced Evaluation

9. What was the *primary focus* of the evaluation:
 - g. Needs assessment to verify and map the extent of a problem (Hanson, 1998, p. 44)
 - h. Formative evaluation to improve and refine the program (Hanson, 1998, p. 44)
 - i. Process evaluation to aid program implementation and replication (Hanson, 1998, p. 44)
 - j. Outcome/impact evaluation to assess program effectiveness in producing change (Hanson, 1998, p. 44)
 - k. Summative evaluation to assess program quality/impact for accountability and decision-making purposes (Hanson, 1998, p. 44)
 - l. Other _____
10. Did the evaluation planning process include determinations of:
 - a. Program evaluability (Jackson, 2004)
 - b. Agreement/contract terms (Jackson, 2004)
 - c. Evaluation focus (Jackson, 2004)
 - d. Methodology (Jackson, 2004)
 - e. Evaluation field test (Jackson, 2004)
 - f. Data collection (Jackson, 2004)
 - g. Data editing and analysis (Jackson, 2004)
 - h. Reporting (Jackson, 2004)
 - i. Other _____

11. Were youth actively involved in the evaluation planning and implementation process (i.e., in roles other than program participant)? (Checkoway & Richards-Schuster, 2003; Sabo, 2003)

12. What challenges to the implementation of the evaluation plan had to be overcome? (Hanson, 1998)

- j. Rigid organizational procedures or settings (Taut & Alkin, 2003; Weiss, 1998a; Cecil, 2004)
- k. Inadequate or problematic data collection, management, analysis, or reporting (Donaldson, Gooler & Scriven, 2002; Taut & Alkin, 2003; Kegeles, Rebchook, & Tebbetts, 2005; Stoecker, 2007)
- l. Disagreement/debate over who should be conducting the evaluation (Stoecker, 2007)
- m. Lack of cooperation, collaboration, or communication among staff (Weiss, 1998a; Donaldson, Gooler, & Scriven, 2002)
- n. Limited methodological capacity of staff (Hasenfeld & English, 1974; Kegeles, Rebchook, & Tebbetts, 2005; Stoecker, 2007)
- o. Lack of staff resources such as time and funding (Gilliam, Davis, Barrington, Lacson, Uhl, & Phoenix, 2002; Hasenfeld & English, 1974; Taut & Alkin, 2003; Kegeles, Rebchook, & Tebbetts, 2005; Stoecker, 2007)
- p. Unexpected/unanticipated evaluation results or consequences (Donaldson, Gooler, & Scriven, 2002; Taut & Alkin, 2003)
- q. Friction between youth organization and funder (Stoecker, 2007; Gilliam, Davis, Barrington, Lacson, Uhl, & Phoenix, 2002)
- r. Misapplication of theories such as constructionism or post-modernism (Weiss, 1998a; Thayer, 2006)
- s. Evaluation results not used (Donaldson, Gooler, & Scriven, 2002)
- t. Other _____

13. What threats to the validity of the evaluation were encountered? (Donaldson, Gooler, & Scriven, 2002; Jackson, n.d.)

- j. Outside forces affected intervention or results (Jackson. n.d; Taut & Alkin, 2003; Weiss, 1998a; Mulroy & Lauber, 2004, Peterson & Randall, 2006)
- k. Interventions mutated from the original design (Jackson. n.d; Peterson & Randall, 2006)
- l. Control groups were not comparable (initially or over time) (Jackson. n.d)
- m. Youth reacted to please or displease program operators (halo/cloud effects) (Jackson. n.d)
- n. Program interventions migrated to the control group (Jackson. n.d)
- o. Youth became sensitized to intervention measures over time (Jackson. n.d)
- p. Maturation of participants clouded intervention results (Jackson. n.d)
- q. The evaluation process caused disruptions that impacted results (Jackson. n.d; Weiss, 1998a; Peterson & Randall, 2006)
- r. Measurements became unstable or changed over time (Jackson. n.d)
- s. Other _____

14. What evidence of successful evaluation implementation and usefulness did you observe?
- f. Improved stakeholder focus, ownership, skills (Cecil, 2004; Gilliam, Davis, Barrington, Lacson, Uhl & Phoenix, 2002, Thayer, 2006)
 - g. Improved organizational capacity, impact, viability, credibility (Cecil, 2004; Thayer, 2006)
 - h. Improved program implementation, results (Thayer, 2006; Gilliam, Davis, Barrington, Lacson, Uhl, & Phoenix, 2002)
 - i. Greater use of evaluation results (Gilliam, Davis, Barrington, Lacson, Uhl, & Phoenix, 2002)
 - j. Other _____
15. With whom were evaluation results shared?
- i. program sponsors and/or funders (Weiss, 1998a)
 - j. program director and staff (Weiss, 1998a)
 - k. program participants (Weiss, 1998a)
 - l. managers of similar programs (Weiss, 1998a)
 - m. policymakers and opinion leaders (Weiss, 1998a)
 - n. social scientists (Weiss, 1998a)
 - o. other evaluators (Weiss, 1998a)
 - p. the general public (Weiss, 1998a)
 - q. Other _____

Section IV. Final Questions about the most recent grant-funded youth program for which you provided services.

16. Please describe a serious implementation challenge or threat to validity that you encountered and successfully dealt with during the above-referenced evaluation. What was the challenge or threat, how did you deal with it, and with what effect?"
17. What evaluation resources (guides, publications, websites, etc.) did you find to be particularly helpful or effective in working with the youth organization?
18. Do you have any other comments about the challenges of evaluating grant-funded programs serving youth ages 14-25 that might be of use to the profession?

Appendix C

Pilot Protocol

Your responses *will* be included in the study results, so please complete the survey “for real.”

1. Read the e-mail inviting you to participate in the study.
 - a. Was your e-mail address the only one shown in the “To” line? Yes/No
 - b. Was the invitation easy to understand? Yes/No
 - c. Were the instructions for responding clear? Yes/No
2. Please indicate any changes to the e-mail invitation that would make people more likely to complete the survey:
3. Click on the link in the e-mail to participate in the study. Complete the entire survey, noting start and end times—but do not click “Done” at the end--yet.

Start time _____ End time _____

Was the time needed to complete the survey reasonable? Yes/No

After recording your start and end times, please answer the following questions about the survey:

4. Was the survey design pleasing? Yes/No
5. Was the font easy to read? Yes/No
6. Did you experience any technical problems with SurveyMonkey? Yes/No
If yes, what were they?
7. Were any of the survey *questions* unclear? Yes/No
If yes, indicate the question number(s):
8. Were *response choices* for any of the survey questions unclear? Yes/No
If yes, indicate the question number(s):
9. Were *response choices* for any questions biased or inappropriate? Yes/No
If so, indicate the question number(s):
10. What are your suggestions for correcting the problem(s) you found above?

Now click "Done" to submit your survey responses to SurveyMonkey. E-mail your responses above to maryp@gwu.edu. If you prefer to discuss them in a conference call, please e-mail me the most convenient dates and times for you.

Thank you!

Appendix D

Summary Results of the Pilot Survey

1. What best describes the TYPE of youth organization you served?

Answer Options	Response Percent	Response Count
Individual non-profit or community-based organization or 501(c)(3)	47.4	9
Public school or school district or component thereof	21.1	4
Four year college or university	10.5	2
Community or junior college	0	0
Government agency	5.3	1
Collaboration of all or some of the above groups	10.5	2
Other (please specify)	5.3	1
- Juvenile justice and status issues (truancy, behavior, etc.)		
TOTAL		19

2. What was the PRIMARY FUNDING SOURCE for the grant program that was evaluated?

Answer Options	Response Percent	Response Count
Federal government	52.6	10
State government	26.3	5
Local or municipal government	5.3	1
Foundation	10.5	2
Corporation	0	0
Other (please specify)	5.3	1
- Unsure; combination of local/state contracts, service fees, foundation grants		
TOTAL		19

3. What type of youth (within the age range 14-25) was the PRIMARY TARGET of the grant program?

Answer Options	Response Percent	Response Count
Academically at-risk youth	26.3	5
Academically gifted youth	0	0
Low-income youth	0	0
Minority youth	15.8	3
English Language Learners	5.3	1
Youth with disabilities	0	0
Gay/lesbian/bisexual/transgender youth	0	0
All youth within the geographic scope of the grant	21.1	4
Other (please specify)	31.6	6
- American Indian youth with homelessness, etc.		
- Runway and homeless (2 responses)		
- Youth with Serious Emotional Disturbance		
- Academically at risk, low-income, minority youth		
- Youth at risk of truancy or ungovernable behavior		
TOTAL		19

4. What type of youth outcome was the PRIMARY FOCUS of the grant program?

Answer Options	Response Percent	Response Count
High school completion or graduation	26.3	5
Career development or employment	5.3	1
College enrollment or completion	0	0
Independent living	0	0
Healthy choices	21.1	4
Other (please specify)	47.4	9
- Shelter, preventing out-of-home placements (2 responses)		
- Teacher training/education (2 responses)		
- Healthy transition to adulthood		
- Mental health services		
- Engagement in scientific research		
- Improved reading and math skills		
- All of the above		
TOTAL		19

5. What was the PRIMARY REASON that the youth organization hired an evaluator or evaluation consultant?

Answer Options	Response Percent	Response Count
Lack of organizational capacity to conduct an evaluation	5.3	1
Funder requirement	68.4	13
External requirement such as accreditation	0	0
Stakeholder decision	5.3	1
Board of Director or senior management decision	5.3	1
Other (please specify)	15.8	3
- Funder requirement & senior management decision		
- Funder reporting purposes & program improvement		
- Lack of capacity		
TOTAL		19

6. For what type of organization did you work when providing evaluation services to the youth organization?

Answer Options	Response Percent	Response Count
The youth organization being evaluated	10.5	2
Yourself (i.e., self-employed consultant)	26.3	5
Independent consulting firm	10.5	2
College or university	31.6	6
Government agency	5.3	1
Corporate or private foundation	0	0
Other (please specify)	15.8	3
- Non-profit company/corporation/service provider (3 responses)		
TOTAL		19

7. What BEST DESCRIBES your role in providing evaluation services to the youth organization?

Answer Options	Response Percent	Response Count
Judge of program merit or worth	15.8	3
Auditor, inspector, or compliance/accountability evaluator	10.5	2
Researcher or scientist	5.3	1
Evaluation facilitator or guide	31.6	6
Team member with evaluation perspective or internal advisor	5.3	1
Collaborator or partner	0	0
Program improvement consultant	0	0
Empowerment facilitator	0	0
Advocate for social change or support of a cause	5.3	1
Educator or capacity builder	5.3	1
Organization development consultant or change agent	0	0
Program advocate	0	0
Critical friend	5.3	1
Other (please specify)	15.8	3
- Judge of merit or worth & organizational development consultant or change agent		
- Internal evaluator		
- Focus group moderator for evaluation		
	TOTAL	19

8. Did you encounter any staff and/or stakeholder ANXIETY related to their:

Answer Options	None at all	A slight amount	A small amount	A moderate amount	A great deal
Lack of experience/knowledge with evaluation	2	7	6	3	1
Negative past experience with evaluation	5	6	3	2	3
Fear of failure, negative results, or consequences	3	4	3	4	5
Organizational culture or social norms resistant to evaluation	4	2	6	6	1
Role conflicts between evaluator(s) and stakeholders	8	4	5	1	1
Concerns about evaluator competence (technical or social)	14	3	1	1	0
Concerns about evaluator understanding of program environment	7	9	0	2	1
Other (please specify)					
- A moderate amount of discussion about me being an "outsider" to the ethnicity being served					
				TOTAL	19

9. What was the PRIMARY FOCUS of the evaluation?

Answer Options	Response Percent	Response Count
Needs assessment to verify and map the extent of a problem	0	0
Formative evaluation to improve and refine the program	21.1	4
Process evaluation to aid program implementation and replication	10.5	2
Outcome/impact evaluation to assess program effectiveness in producing change	47.4	9
Summative evaluation to assess program quality/impact for accountability and decision-making purposes	10.5	2
Other (please specify)	10.5	2
- Outcome/impact & formative		
- Process & outcome		
TOTAL		19

10. Did the evaluation planning process include determinations of:

Answer Options	No	Yes
Program evaluability	8	11
Agreement/contract terms	3	16
Evaluation focus	2	17
Methodology	2	17
Evaluation field test	14	5
Data collection	1	18
Data editing/reporting	1	18
Reporting	2	17
Other (please specify)		
- Replicated the same evaluation plan that had been in use for years		
- Came in midway through the process/not involved in planning		
TOTAL		19

11. Were youth actively involved in the planning and implementation process, i.e., in roles other than program participant?

Answer Options	None at all	A slight amount	A small amount	A moderate amount	A great deal
Youth were actively involved	10	4	2	1	2

13. What threats to the validity of the evaluation were encountered? Please select N/A only for the last two options and only if you did not use control groups.

Answer Options	None at all	A slight amount	A small amount	A moderate amount	A great deal	N/A
Outside forces affected intervention or results	7	2	4	3	2	1
Interventions mutated from the original design	6	5	2	3	1	2
Youth reacted to please or displeasure program operator (halo/cloud effects)	8	5	3	1	0	2
Youth became sensitized to intervention measures over time	9	4	2	1	0	3
Maturation of youth participants biased the observed results	13	0	1	2	1	2
Evaluation process caused disruptions that impacted results	12	2	2	1	0	2
Measurements became unstable or changed over time	10	4	2	1	0	2
Control groups were not comparable (initially or over time)	1	2	1	0	0	15
Program interventions migrated to the control groups	4	1	0	0	0	14
Other (please specify)						
- Ours was not an intervention evaluation						
TOTAL	19					

14. What evidence of successful evaluation implementation and usefulness were perhaps observed?

Answer Options	None at all	A slight amount	A small amount	A moderate amount	A great deal
Improved stakeholder focus, ownership, and/or skills	2	3	6	5	2
Improve organizational capacity, impact, viability, and/or credibility	3	3	6	4	3
Improved program implementation	2	3	6	3	3
Improved program results	3	2	6	3	3
Improved use of evaluation results by youth program, funder, or other audience	3	2	6	2	5
Other (please specify)					
- Better understanding of target population					
- Too soon to answer most of these questions					
TOTAL 19					

15. With whom were evaluation results shared?

Answer Options	No	Yes
Program sponsors and/or funders	1	18
Program director and staff	1	18
Program participants	11	7
Practitioners in the youth service field	11	8
Policymakers and/or opinion leaders	6	13
Social scientists	17	1
Other evaluators	13	6
The general public	10	8
Other (please specify)		
- Not yet shared with latter groups, but will be		
- I was not involved in reporting		
- Too early in program/evaluation to have results		
TOTAL 19 for items 1, 2, 4, 5, & 7		
TOTAL 18 for items 3, 6 & 8		

16. Please describe a serious implementation challenge or threat to validity that you encountered and successfully dealt with during the afore-mentioned evaluation. What was the challenge or threat, how did you deal with it, and with what effect? (TOTAL 17)

Three respondents reported no implementation challenges or threats to validity. 14 respondents described problems with staff attitudes/cooperation (7), data (5), measurement (3), time (2), infrastructure, outside forces (unexpected funding increase), stakeholder conflict, and maintaining student anonymity. Five respondents reported more than one problem, which in one case was a ripple effect caused by lack of staff cooperation that resulted in measurement problems and incomplete data. One of the issues with data was “some concern that the numbers collected ... do not reflect the true worth of the program.”

Seven respondents reported positive resolutions to the problems. Six did not say if the problem(s) had been resolved. One reported no resolution to a problem with “extreme personal differences among stakeholders that cause meetings to sometimes deteriorate into...shouting matches.”

17. What evaluation resources (guides, publications, Web sites, etc.) did you find to be particularly helpful or effective in working with the youth organization? (TOTAL 17)

Three people could not think of any resources or reported none. One person said he or she was aware of unspecified resources on youth involvement but that it was not feasible due to logistics and other factors.

Four people reported professional networking or involving another colleague (EVALTALK, colleagues, field networking).

Three people gave non-specific responses (everything, books and journal articles, federal agency guidelines).

Specific publications included Patton’s *Utilization Focused Evaluation*, the Search Institute’s *40 Assets for Youth Development, Getting to Outcomes*, and the *GAO Yellow Book*.

Web-based resources included a wiki, Fetterman’s web site on empowerment evaluation, and the HFRP (Harvard Family Research Project) web site

Other resources included “EDDRA” (which the researcher hopes is the Exchange on Drug Demand Reduction Action and not the Education Disinformation Detection and

Reporting Center), Arizona State University Evaluation Center, and “familiarity with the organization and the community it served.”

18. Do you have any other comments about the challenge of evaluating grant-funded programs serving youth ages 14-25 that might be of use to the profession? (TOTAL 17)

Seven respondents had no other comments.

Other respondents mentioned challenges with

- providing an evidence- based evaluability foundation for agencies
- convincing school districts that are under pressure academically and financially that evaluation is needed
- high transiency of the targeted youth population that causes missing data
- IRB restrictions on direct contact with students
- logistical and other problems with youth engagement and involvement
- getting leadership to support the evaluation and not be afraid of results
- the need for familiarity with the community and population served
- frequent turnover among program staff
- a disconnect between what the funders/program developers expect and what the staff think they are supposed to do
- budgets that do not adequately provide for evaluation services
- the time needed to effect change, including inadequate one-shot trainings and staffers for whom time is in short supply (these issues were addressed by extending the contract)

One respondent reiterated a previous concern about “whether negative or positive changes in the evaluation outcomes actually reflects [sic] changes in student outcomes or overall program quality” because the “strengths” of the consortium members may not be captured by “the quantitative information required by the government.”

Appendix E

Question 17 Responses

17. What evaluation resources (guides, publications, Web sites, etc.) do you find to be particularly helpful or effective in working with youth organizations (not just the youth program described previously)?

AEA – 53 References

Nonspecific AEA Resources- 12 References

American Evaluation Association (AEA); AEA; AEA; AEA; AEA; AEA; AEA; AEA; AEA is a great resource; AEA; AEA; AEA;
AEA memberships

AEA Publications/Materials/Journals - 31 References

*AEA publication dealing specifically with evaluating youth organizations

*AEA publication on evaluation of youth serving organizations.

*AEA journal articles that reference youth programs

AEA Publications; AEA publications; AEA publications (these aren't specific to youth, however); AEA publications; AEA journal; AEA journal; AJE; American Journal of Evaluation; AEA journals; AEA Journal and New Directions publications; AEA Journal;

AEA publications AJE & NDE; New Directions for Evaluation publication; Some articles in AEA journal have been helpful; American Journal of Evaluation; AEA journals; AEA publications; AEA Journal; AEA materials; Journal--New Directions in Evaluation; j of the AEA; journals (New Directions in Evaluation, American Journal of Evaluation);

The American Evaluation Association has had some helpful materials but for the kind of work we're embarking on, I haven't found a lot of information about how to evaluate such systemic initiatives.

I always refer to the Guiding Principles

AEA Guiding Principles are shared with all clients..

American Evaluation Association (Cultural/Contextual Evaluation

AEA Website/listserv – 6 References

American Evaluation Association Website; American Evaluation Association website; AEA website; AEA websites

AEA Listserv -

EVALTALK

Evaltalk listserv, Eval talk listserv; Evaltalk, evaltalk. I don't do enough youth-specific work to identify youth-specific resources, but EVALtalk is always helpful to me; EvalTalk mostly and the resources it provides, AEA listserv; AEA listservs;

Other AEA or AEA-related Resources – 4 References

AEA conferences, AEA meetings

Our local affiliate

The environmental evaluation TIG has had some great session at the annual conference

The Joint Committee on Standards for Education Evaluation – 8 References

Joint committee evaluation standards proved very helpful; Standards for program evaluation, Eval Center, WMU;
 The Standards for Program Evaluations (Joint Committee) were especially helpful in planning the evaluation; program evaluation standards; Evaluation Standards; Joint Committee Standards for Educational Evaluations; I always refer to the Program Evaluation Standards; The Program Evaluation Standards; AEA evaluation standards

Reference Books/Texts - 27 References

Michael Quinn Patton – 14 references

Patton, Patton, Patton's work in evaluation, Patton's Utilization-Focused Evaluation, Utilization-focused eval book, M.Q. Patton publications, Utilization-focused Evaluation (MQ Patton), Patton- utilization focused evaluation, Patton's "Utilization-Focused Evaluation," Michael Pattons publications and books, Patton's Utilization Focused Evaluation; , Patton books on Utilization Focused Evaluation, Patton; Patton,

Rossi, Freeman & Lipsey – 4 references

I use the Rossi text usually for all evaluations; Evaluation by Rossi, Freeman and Lipsey; We used Rossi, Lipsey and Freeman publication to help determine formulate our evaluation protocol; Rossi et al

Greene – 2 references

Greene; A text by Jennifer Greene was very influential.

Stakes/Stake – 2 references

Wholey et al – 2 references

Wholey et al's Handbook of Practical Program Evaluation, Wholey, Hatry and Newcomer's text;

Miles and Huberman's sourcebook for quaitative data analysis

Dr. Liliana Rodriguez-Campos's book on collaborative evaluations

Sechrest

Wandersman

"Confirmative Evaluation" (2004) by Joan C. Dessinger and James L. Moseley.

Poister

Brownson et al

Fink

Rist

Cronbach,

Issel text

Weiss

Chen, H. T. (2005), Practical program evaluation. Sage Publications. Thousand Oaks, CA.

We used the literature, but especially Tom Schwandt's "Evaluation Practice Reconsidered" (Peter Lang, 2002)

Jacqueline Ferguson's workshop & publications

Rea & Parker's Designing & Conducting Survey Research

The Festin and Philbin 2007 text has been helpful
 The newer Sage booklet, Nonprofit Evaluation
 *Kim Sabo Flores
 The Russell Sage box set on focus groups is especially helpful.
 Social Experiments by Larry Orr;
 *McLaughlin, M. (2000). Community counts: How youth organizations matter for youth development. Washington, D.C. Public Education Network
 Russ-Eft & Preskill book.

Journals/Research – 14 References

Educational Researcher – 2 references
 ERIC – 2 references
 Evaluation Review
 Evaluation journals
 Researched some on the content of journals as we were using them as one aspect of data collection
 Journal of Community Psychology.
 The Blueprint studies [drug prevention studies?].
 Evaluation & Program Planning and recognized evaluators
 Scientific Journals such as Accident Analysis and Prevention
 Journal of Studies on Alcohol,
 I read various journals, including EEPA, EP. These keep me up on things.
 Journal of Educational Research
 AERA journals

Other Publications – 35 References

Publications Related to Youth – 17 References

- * The Colorado Trust toolkit for Evaluating Positive Youth Development
- *Child Trends publications, such as: Bronte-Tinkew, J., Moore, K. A., & Shwalb, R. (2006). Measuring outcomes for children and youth in out-of-school time programs: Moving beyond measuring academics. Washington, DC: Child Trends.
- *National Center on Secondary Education and Transition E-News
- *Measuring the Quality of Mentor-Youth Relationships (August 2002) report by the Public/Private Ventures (P/PV)
- *Student Mentoring Programs: Education's Monitoring and Information Sharing Could Be Improved. GAO-04-581, June 25, 2004
- *Research and Training Center for Children's Mental Health
- *youthtoday.org report, "School Based Mentoring - Does it Make the Grade?"
- *Wagner, M., Newman, L., Cameto, R., Levine, P., and Marder, C. (2007). Perceptions and Expectations of Youth With Disabilities. A Special Topic Report From the National Longitudinal Transition Study-2 (NLTS2) (NCSER 2007-3006)
- *Forum for Youth Investment compilations and critiques
- * The National Child Traumatic Stress Network publications
- *The Search Institute Asset-Based variables

*National Mentoring Center, "Building Relationships. A Guide for New Mentors." April 2001

* Search Institute literature on capacity building and resiliency --there is also a tool kit on measures for youth outcome -but can't think what it is called at the moment.

*The Hawkins and Catalano article on youth risk and protective factors was most helpful.

*Lerner's Positive Youth Development literature

*National Research Center Youth Outcome toolkit

United Way publications – 3 References

Getting to Outcomes; United Way has a very nice primer on evaluation of organizations that I used.

I used the Measuring Program Outcomes: A Practical Approach published by United Way of America as the core text for the plans.

Community Anti-Drug Coalitions of America (CADCA) – 3 references

Community Anti-Drug Coalitions of America (CADCA) Strategic Planning Framework Primers; This was a substance abuse prevention program; resources provided by CADCA particularly helpful; CADCA.

Other Publications – 11 references

Univ. of Wisconsin Cooperative Extension Planning a Program Evaluation Basic Guide to Outcomes-Based Evaluation for Nonprofit Organizations with Very Limited Resources

*Harvard Evaluation Exchange Newsletter has lots of good resources focused on education, family interventions etc.

NACADA Advising assessment guidelines

PND (Philanthropy News Digest) Publication Alerts

National Implementation Research Network monograph

ILO-OCED publications

*Education Week

Evaluation Exchange

*Learning Point Associates' Beyond The Bell toolkit

Society for Prevention Research Community Toolkit

Univ of Kansas Community Tool Box

Government Resources – 51 References

Federal resources – 10 references

TA (Technical Assistance); websites for particular program; US government, federal websites,

Federal Evaluation Guidelines, Guides available from federal government to how to do evaluation of this program, Federal guides/templates on program evaluation. Federal guidelines for the program, Federal regulations, rfp, etc., federal program meetings.

Resources produced through the federal programs that provided these "flow-through" funds to the state (e.g. 21st Century Community Learning Center funds).

Centers for Disease Control and Prevention (CDC) – 7 references

CDC evaluation website, CDC DASH evaluation website; CDC; The Centers for Disease Control and Prevention; YRBS, NHANES III, and SLAITS data from CDC; the CDC <http://www.cdc.gov/ncipc/pub-res/measure.htm>; CDC resources; CDC HIV surveys: <http://www.caps.ucsf.edu/tools/surveys/>

CFYAR (USDA/CSREES) tool kits on evaluation

Dept. of Ed. Websites – 2 references

Federal DOE sites providing FAQ or clearinghouse information .

Department of Justice – 9 references

DOJ evaluation resources; DOJ website

Office of Juvenile Justice and Delinquency Prevention (OJJDP) publications, OJJDP, US Office of Juvenile Justice

NIC (National Institute of Corrections),

BJA (Bureau of Justice Assistance), BJA Center for Program Evaluation

NIJ (National Institute of Justice) publications resources

National Institutes of Health (NIH)

National Science Foundation (NSF) – 4 references

NSF evaluation handbooks; NSF program evaluation guide; Online Evaluation Resource Library (OERL) [of the National Science Foundation]

Department of Health and Human Services – 15 references

Dept HHS Admin for Children and Families websites; USH&HS workshops

Substance Abuse and Mental Health Services Administration (SAMHSA);

SAMHSA resources; On-line SAMHSA publications; SAMHSA; SAMHSA;

Have used SAMSHA web site and link; SAMHSA website and documentation;

SAMSHA's website; SAMHSA Evidence-based practices/materials; SAMHSA

eNetwork, SAMHSA websites related to the project in other communities;

SAMHSA website; Prevention platform

Department of Labor, Office of Disability Employment Policy (ODEP) website

I found at least a dozen such on the website *.gov [www.usa.gov] in the new abstinence webpage

State DOE websites

Foundation Resources – 16 References

W.K. Kellogg Foundation - 13 references

logic model, logic model guide, guides on cluster evaluation, handbook/materials, website & publications

We still find the Kellogg Foundation's guides to logic modeling to be helpful in focusing the evaluation and helping programs explain the underlying structure and theory of action of their programs.

Annie E. Casey Foundation – 2 references

Kresge Foundation Evaluation model

Gates Foundation

Web Sites/Internet – 41 References

*Harvard websites – 5 references

harvard website; Harvard Family Research Project www.hfrp.org; www.hfrp.org database and summaries; Harvard Family Research Project website; The website of the Harvard Family Research Project

*Public/Private Ventures – 2 references

www.ppv.org; Many reports on mentoring available on the P/PV website.

*Youth development models such as that articulated by Gambone & Connell at Youth Development Strategies, Inc., www.ydsi.org

*The website of the Forum for Youth Investment.

*National Collaborative on Workforce and Disability for Youth (<http://www.ncwd-youth.info/>)

*The National Child Traumatic Stress Network web site

*SEDL toolkit (Southwest Educational [Development] Laboratory) website

*The National Wraparound Initiative website is very useful

*Collaborative for Academic Social and Emotional Learning (CASEL) website

AERA – 2 references

AERA website, We have used a variety of websites and publications in our evaluations, but one we have consistently found to be useful of late has been the American Education Research Assn (AERA) site.

ERIC – 2 references

<http://eric.ed.gov/>;

http://eric.ed.gov/ERICWebPortal/resources/html/about/about_eric.html

Western Michigan evaluation center website – 2 references

APHA website

ASHA website

CDC website

Carter McNamara's webpages

Dr. Liliana Rodriguez-Campos's website

Penn State Cooperative Extension Evaluation webpages

Idealist.com

ILO-OCED websites

ETR Associates website

UCLA mental health web site;

Research Methods Knowledge Base - <http://www.socialresearchmethods.net/kb/>

Univ of Wisconsin Extension website

<http://www.spss.com/PDFs/STIPlr.pdf>

<http://www.eval.org/Resources/Collections.asp>

http://edweb.sdsu.edu/courses/ed690sp05/SurveyDesign_ResponseOption.pdf

<http://www.ryerson.ca/~mjoppe/ResearchProcess/WriteBetterQuestion.htm>

<http://survey.pearsonncs.com/planning/response-scales.htm>

<http://www.myonlineforms.com/surveyguru/index.php3?pg=construction>

<http://gearupdata.org/>

www.eoutcome.org

www.nbowmanconsulting.com (Indigenous Evaluation)

Google Scholar

Peer/Stakeholder Networks/Communications – 25 Resources

AERA memberships and listservs

Communications with peers.

Personally asking someone who has A LOT of experience conducting evaluations and dealing with the variety of problems that arise

(Stufflebeam, Sanders, etc.). This applies to any problems regardless of whether it is a youth organization or not. In the field of education the majority of organizations are youth organizations.

The most important evaluation resource that was particularly helpful and effective with youth grantees were/are the evaluator's staff

(people resource). Evaluator staff worked with youth organizations' staff, providing them with ongoing support, training, and evaluation problem-solving. We provide written resources, access to tools and information online, and ensure youth staff know who to contact with questions/clarifications and how to contact those staffers. Because grantees are direct service providers, they appreciated the proactive approach.

Curriculum developers (with whom I work),

Consultants on the technical assistance team

Focus groups

Stakeholder group feedback

Feedback from clients

The study was part of a larger initiative, so peer support and collaboration was very helpful, as well as consultation with the outside evaluator.

Sharing resources with peer evaluators who are also working on evaluation of youth programs in my field.

Talking with colleagues.

Colleagues in private practice and at universities.

Professional networks that include people who have both a history of working with and a perceived credibility with the youth program to ensure greater participation.

informal knowledge sharing with co-workers.

direct service providers and their supervisors [&] participants

Evaluator's Institute colleagues

I also keep up with providers and their ongoing evaluations, which let me know what others are doing and how.

Collaboration on the question to be asked.

Information/reports about previous evaluations conducted on programs operated/implemented by peer organizations

Talking with other experts in the field

I find that being there and observing prior to evaluation activities is more useful than any resources, showing youth your investment in them, rather than outcomes.

Use of a professional youth worker organization as a partner in recruiting programs for the study.

I have many resources, but I rely most on the experts with whom I work. The evaluations are true collaborations between my expertise

and experience in R&E and theirs in their direct service programs.

Instruments/Data/Software/Technology – 23 References

Instruments – 9 references

Catalogues of pre-existing instruments; Other bullying evaluation instruments; Articles and tech reports on psychometrics of various scales used to assess risk/protective factors; Validated measurement tools were already in place (although we discovered numerous problems with the tools); Developmental Assets Profile (a measure) from the Search Institute is a useful tool, since it has credibility with funders; CDC's Youth Risk Behavior Survey; Rhodes & Grossman youth survey; YPQA

Data – 10 references

NC College Tech Prep Longitudinal Data; School data, PEIMS, AEIS, JUCR, City County generated GIS Reports, Student Survey data; epidemiologic data; Census demographics; We need benchmark data for almost all the data we produce, so any website containing comparative data is useful.

Software – 2 references

INVIVO in our data analyses; The internet-based free statistical packages such as the one maintained at Vassar College Psychology Department; Good software for both qualitative and quantitative data analysis.

Using text messaging to reach youth who no longer check emails regularly

I.T in general, and in particular, texting and internet are rich resources for communicating with youth organizations.

Experience/Training/Professional Development/Education – 15 References

I do not typically use resources beyond my personal evaluation experience and ongoing education.

Evaluators' Institutes;

Evaluator's Institute trainings & class materials

Previous training,

My own experience.

I have a background in youth development work, was drawing on my own areas of expertise.

Course taught by Jennifer Greene at the University of Illinois

Conferences

personal experience

I've used some resources from the NAAEE annual conference - mostly information presented during conferences.

My experience during the 7 years I spent at Harvard's Graduate School of Education conducting a large national evaluation of a

professional development intervention provided the foundation I needed for the evaluation described in this survey. (Of course, having taught middle school and high school science, and created schools for dropouts for the 25 years previous to my doctoral work, provided the real foundation on which my Harvard evaluation experience was built.)

We used our past experience and suggestions by the funding agency.

AERA meetings

I haven't used any sources that were particularly designed for youth organizations. I have 25 years of experience working with youth directly.

Evaluator/Stakeholder Developed Resources/Procedures – 10 References

One of the stakeholders in the evaluation pointed us to materials that it had developed pertaining to conducting research with

community-based organizations (CBOs). These materials helped to highlight the perspective of the CBO staff and why they might be resistant to evaluation.

I usually rely on materials generated in previous projects I have been associated with, in addition to conducting periodic web searches

and searches of journal databases.

There is no one resource I use all the time; it varies from case to case. I have been fortunate to work with program directors who have

strong evaluation focus, so I haven't had to do much. I produce most of my own materials in working with program line-staff.

We created our own toolkits and program manual. Integrating evaluation and assessment directly into the manual helped to make it

more effective. At the time we started, there were not the resources available that there are now (haven't done this work since 2004).

I have developed a framework particularly useful for working with youth. Based on Self-Determination Theory, it qualifies

achievement with well-being, measures the effect of evaluation feedback on both staff and youth participants, and measures how well both participants and staff feel supported by various aspects of the program system.

created internal evaluation training resources (online video) and an online evaluation toolkit

Capacity building on an individual basis with each program director/staff. These short sessions to review progress, answer questions

or sort out small details are invaluable. We have two guides: one a down and dirty "how to" for a quick desk references with FAQs and process explanations as well as a focus group guide. We adapted materials that were readily available on line from the Innovation Network, GTO (Wandersman, Harvard eval and my own online materials. We set up an interactive (primitive, but interactive network) as well. The best, however, were the one to one sessions either on the phone or in person. The phone worked best since scheduling, especially for something like evaluation, is difficult. The organization and stakeholders want to "do it right" and honestly try to as they are all highly involved in the process. The evaluator, IMHO, must realize that most longer term evaluations require what can be overwhelming amounts of money, time and trust on the front end if one expects valid organization-driven outcomes on the back end.

My own developed over a 40-year career in evaluation

We do a lot of upfront work with groups to place evaluation in a larger context of program management. Spend time helping them use

logic models as a tool

We translate standard evaluation procedures (e.g., logic models and stage matrixes such as CIPP) into our own materials, which we present during the planning of or application for evaluation contracts.

I teach evaluation, so I used materials from my classes, including several journal articles, some power point presentations in particular.

Other Resources - 43

*Harvard – 7 references

Harvard Out-of-School Time Program Evaluation; HFRP has great resources for afterschool programs; Harvard Family Research Project, The Harvard Family Research Project's many resources relating to Out-of-School-Time program research were very helpful in identifying appropriate measures for these populations, HFRP, Harvard afterschool web-site; Harvard Family Research Project

Evaluation Center at Western Michigan University – 3 references

Western Michigan University, CIPP (Context, Inputs, Processes, Products) model; evaluation center at WMU; West Michigan

University of Wisconsin Extension – 3 references

Wisconsin Extension Service; U Wisconsin; University of Wisconsin-Extension resources

*Search Institute – 3 references

United Way – 2 references

*4H

*Adelman and Taylor (UCLA);

*Afterschool Alliance has great resources for afterschool programs.

American Psychological Association (APA)

Beth Miller's work

*Chicago Consortium for School Research

*Cooperative Extension Information on youth development and other organizations

David Morgan's work on focus groups

Endowment for Health

Everett Rogers

*High Scope

*Institute for Higher Education Policy

MDRC

*National Association of Student Assistance Programs

*National Campaign to Prevent Teen Pregnancy

*National Collaboration for Youth

*Pathways to College Network

*PEPNet

*Permanency Planning resources.

Policy associates OST reports

Policy Studies Associates

*Portland Research and Training Center for Family Support and Children's Mental Health

Public Health Agency of Canada

Urban Institute

Nonspecific Resources – 77

E.g., “books,” “websites,” “journals,” “resources specific to arts-based youth program evaluations”...

Not Sure What Was Meant – 3 References

NREP website, ISE recommendations on evaluation; AOI

Everything/Many/Various/Depends – 7 Responses

None/Nothing/Don’t Know/Not Sure/Can’t Recall – 48 Responses

Non-Responses - 58

Comments - 4