Evaluation Checklist Project

Qualitative Evaluation Checklist

Michael Quinn Patton

The purposes of this checklist are to guide evaluators in determining when qualitative methods are appropriate for an evaluative inquiry and factors to consider (1) to select qualitative approaches that are particularly appropriate for a given evaluation's expected uses and answer the evaluation's questions, (2) to collect high quality and credible qualitative evaluation data, and (3) to analyze and report qualitative

1. Determine the extent to which qualitative methods are appropriate given the evaluation's purposes and intended uses.
2. Determine which general strategic themes of qualitative inquiry will guide the evaluation. Determine qualitative design strategies, data collection options, and analysis approaches based on the evaluation's purpose.
3. Determine which qualitative evaluation applications are especially appropriate given the evaluation's purpose and priorities.
4. Make major design decisions so that the design answers important evaluation questions for intended users. Consider design options and choose those most appropriate for the evaluation's purposes.
5. Where fieldwork is part of the evaluation, determine how to approach the fieldwork.
6. Where open-ended interviewing is part of the evaluation, determine how to approach the interviews.
7. Design the evaluation with careful attention to ethical issues.
8. Anticipate analysis—design the evaluation data collection to facilitate analysis.
9. Analyze the data so that the qualitative findings are clear, credible, and address the relevant and priority evaluation questions and issues.
10. Focus the qualitative evaluation report.

Introduction

Qualitative evaluations use qualitative and naturalistic methods, sometimes alone, but often in combination with quantitative data. Qualitative methods include three kinds of data collection: (1) in-depth, open-ended interviews; (2) direct observation; and (3) written documents.

Interviews: Open-ended questions and probes yield in-depth responses about people's experiences, perceptions, opinions, feelings, and knowledge. Data consist of verbatim quotations with sufficient context to be interpretable.

Observations: Fieldwork descriptions of activities, behaviors, actions, conversations, interpersonal interactions, organizational or community processes, or any other aspect of observable human experience. Data consist of field notes: rich, detailed descriptions, including the context within which the observations were made.

Documents: Written materials and other documents from organizational, clinical, or program records; memoranda and correspondence; official publications and reports; personal diaries, letters, artistic works, photographs, and memorabilia; and written responses to open-ended surveys. Data consist of excerpts from documents captured in a way that records and preserves context.

The data for qualitative evaluation typically come from fieldwork. The evaluator spends time in the setting under study—a program, organization, or community where change efforts can be observed, people interviewed, and documents analyzed. The evaluator makes firsthand observations of activities and interactions, sometimes engaging personally in those activities as a "participant observer." For example, an evaluator might participate in all or part of the program under study, participating as a regular program member, client, or student. The qualitative evaluator talks with people about their experiences and perceptions. More formal individual or group interviews may be conducted. Relevant records and documents are examined. Extensive field notes are collected through these observations, interviews, and document reviews. The voluminous raw data in these field notes are organized into readable narrative descriptions with major themes, categories, and illustrative case examples extracted through content analysis. The themes, patterns, understandings, and insights that emerge from evaluation fieldwork and subsequent analysis are the fruit of qualitative inquiry.

Qualitative findings may be presented alone or in combination with quantitative data. At the simplest level, a questionnaire or interview that asks both fixed-choice (closed) questions and open-ended questions is an example of how quantitative measurement and qualitative inquiry are often combined.

The quality of qualitative data depends to a great extent on the methodological skill, sensitivity, and integrity of the evaluator. Systematic and rigorous observation involves far more than just being present and looking around. Skillful interviewing involves much more than just asking questions. Content analysis requires considerably more than just reading to see what's there. Generating useful and credible qualitative findings through observation, interviewing, and content analysis requires discipline, knowledge, training, practice, creativity, and hard work.

Qualitative methods are often used in evaluations because they tell the *program's story* by capturing and communicating the *participants' stories. Evaluation* case studies have all the elements of a good story. They tell what happened when, to whom, and with what consequences. The purpose of such studies is to gather information and generate findings that are useful. Understanding the program's and participant's stories is useful to the extent that those stories illuminate the processes and outcomes of the program for those who must make decisions about the program. The methodological implication of this criterion is

that the intended users must value the findings and find them credible. They must be interested in the stories, experiences, and perceptions of program participants beyond simply knowing how many came into the program, how many completed it, and how many did what afterwards. Qualitative findings in evaluation can illuminate the people behind the numbers and put faces on the statistics to deepen understanding.

1. Determine the extent to which qualitative methods are appropriate given the evaluation's purposes and intended uses.

- ☐ Be prepared to explain the variations, strengths, and weaknesses of qualitative evaluations.
- ☐ Determine the criteria by which the quality of the evaluation will be judged.
- ☐ Determine the extent to which qualitative evaluation will be accepted or controversial given the evaluation's purpose, users, and audiences.
- ☐ Determine what foundation should be laid to assure that the findings of a qualitative evaluation will be credible.

2. Determine which general strategic themes of qualitative inquiry will guide the evaluation. Determine qualitative design strategies, data collection options, and analysis approaches based on the evaluation's purpose.

- ☐ *Naturalistic inquiry*: Determine the degree to which it is possible and desirable to study the program as it unfolds naturally and openly, that is, without a predetermined focus or preordinate categories of analysis.
- ☐ *Emergent design flexibility:* Determine the extent to which it will be possible to adapt the evaluation design and add additional elements of data collection as understanding deepens and as the evaluation unfolds. (Some evaluators and/or evaluation funders want to know in advance exactly what data will be collected from whom in what time frame; other designs are more open and emergent.)
- ☐ *Purposeful sampling:* Determine what purposeful sampling strategy (or strategies) will be used for the evaluation. Pick cases for study (e.g., program participants, staff, organizations, communities, cultures, events, critical incidences) that are "information rich" and illuminative, that is, that will provide appropriate data given the evaluation's purpose. (Sampling is aimed at generating insights into key evaluation issues and program effectiveness, not empirical generalization from a sample to a population. Specific purposeful sampling options are listed later in this checklist.)
- ☐ *Focus on priorities:* Determine what elements or aspects of program processes and outcomes will be studied qualitatively in the evaluation.
    - Decide what evaluation questions lend themselves to qualitative inquiry, for example, questions concerning what outcomes *mean* to participants rather than how much of an outcome was attained.

- Determine what program observations will yield detailed, thick descriptions that illuminate evaluation questions.
- Determine what interviews will be needed to capture participants' perspectives and experiences.
- Identify documents that will be reviewed and analyzed.

☐ *Holistic perspective:* Determine the extent to which the final evaluation report will describe and examine the *whole* program being evaluated.

- Decide if the purpose is to understand the program as a complex system that is more than the sum of its parts.
- Decide how important it will be to capture and examine complex interdependencies and system dynamics that cannot meaningfully be portrayed through a few discrete variables and linear, cause-effect relationships.
- Determine how important it will be to place findings in a social, historical, and temporal context.
- Determine what comparisons will be made or if the program will be evaluated as a case unto itself.

☐ *Voice and perspective*: Determine what perspective the qualitative evaluator will bring to the evaluation.

- Determine what evaluator stance will be credible. How will the evaluator conduct fieldwork and interviews and analyze data in a way that conveys authenticity and trustworthiness?
- Determine how balance will be achieved and communicated given the qualitative nature of the evaluation and concerns about perspective that often accompany qualitative inquiry.

3. Determine which qualitative evaluation applications are especially appropriate given the evaluation's purpose and priorities.

Below are evaluation issues for which qualitative methods can be especially appropriate. This is not an exhaustive list, but is meant to suggest possibilities. The point is to assure the appropriateness of qualitative methods for an evaluation.

☐ *Checklist of standard qualitative evaluation applications*—determine how important it is to:

- Evaluate individualized outcomes—qualitative data are especially useful where different participants are expected to manifest varying outcomes based on their own individual needs and circumstances.
- Document the program's processes—process evaluations examine how the program unfolds and how participants move through the program.
- Conduct an implementation evaluation, that is, look at the extent to which actual implementation matches the original program design and capture implementation variations.

- Evaluate program quality, for example, quality assurance based on case studies.
- Document development over time.
- Investigate system and context changes.
- Look for unanticipated outcomes, side effects, and unexpected consequences in relation to primary program processes, outcomes, and impacts.

☐ *Checklist of qualitative applications that serve special evaluation purposes*—determine how important it is to:

- Personalize and humanize evaluation—to put faces on numbers or make findings easier to relate to for certain audiences.
- Harmonize program and evaluation values; for example, programs that emphasize individualization lend themselves to case studies.
- Capture and communicate stories—in certain program settings a focus on "stories" is less threatening and more friendly than conducting case studies.

☐ *Evaluation models*: The following evaluation models are especially amenable to qualitative methods—determine which you will use.

- Participatory and collaborative evaluations—actively involving program participants and/or staff in the evaluation; qualitative methods are accessible and understandable to nonresearchers.
- Goal-free evaluation—finding out the extent to which program participants' real needs are being met instead of focusing on whether the official stated program goals are being attained.
- Responsive evaluation, constructivist evaluation, and "Fourth Generation Evaluation" (see checklist on constructivist evaluation, a.k.a. Fourth Generation Evaluation).
- Developmental applications: Action research, action learning, reflective practice, and building learning organizations—these are organizational and program development approaches that are especially amenable to qualitative methods.

☐ Utilization-focused evaluation—qualitative evaluations are one option among many (see checklist on utilization-focused evaluation).

4. Make major design decisions so that the design answers important evaluation questions for intended users. Consider design options and choose those most appropriate for the evaluation's purposes.

☐ *Pure or mixed methods design:* Determine whether the evaluation will be purely qualitative or a mixed method design with both qualitative and quantitative data.

☐ *Units of analysis:* No matter what you are studying, always collect data on the lowest level unit of analysis possible; you can aggregate cases later for larger units of analysis. Below are some examples of units of analysis for case studies and comparisons.

- People-focused: individuals; small, informal groups (e.g., friends, gangs); families
- Structure-focused: projects, programs, organizations, units in organizations
- Perspective/worldview-based: People who share a culture; people who share a common experience or perspective (e.g., dropouts, graduates, leaders, parents, Internet listserv participants, survivors, etc.)
- Geography-focused: neighborhoods, villages, cities, farms, states, regions, countries, markets
- Activity-focused: critical incidents, time periods, celebrations, crises, quality assurance violations, events
- Time-based: Particular days, weeks, or months; vacations; Christmas season; rainy season; Ramadan; dry season; full moons; school term; political term of office; election period

(Note: These are not mutually exclusive categories)

☐ *Purposeful sampling strategies:* Select information-rich cases for in-depth study. Strategically and purposefully select specific types and numbers of cases appropriate to the evaluation's purposes and resources. Options include:

- Extreme or deviant case (outlier) sampling: Learn from unusual or outlier program participants of interest, e.g., outstanding successes/notable failures; top of the class/dropouts; exotic events; crises.
- Intensity sampling: Information-rich cases manifest the phenomenon intensely, but not extremely, e.g., good students/poor students; above average/below average.
- Maximum variation sampling: Purposefully pick a wide range of cases to get variation on dimensions of interest. Document uniquenesses or variations that have emerged in adapting to different conditions; identify important common patterns that cut across variations (cut through the noise of variation).
- Homogeneous sampling: Focus; reduce variation; simplify analysis; facilitate group interviewing.
- Typical case sampling: Illustrate or highlight what is typical, normal, average.
- Critical case sampling: Permits *logical* generalization and maximum application of information to other cases because if it's true of this one case, it's likely to be true of all other cases.
- Snowball or chain: Identify cases of interest from sampling people who know people who know people who know what cases are information-rich, i.e., good examples for study, good interview subjects.
- Criterion sampling: Pick all cases that meet some criterion, e.g., all children abused in a treatment facility; quality assurance.
- Theory-based or operational construct sampling: Find manifestations of a theoretical construct of interest so as to elaborate and examine the construct and its variations, used in relation to program theory or logic model.

- Stratified purposeful sampling: Illustrate characteristics of particular subgroups of interest; facilitate comparisons.
- Opportunistic or emergent sampling: Follow new leads during fieldwork; taking advantage of the unexpected; flexibility.
- Random purposeful sampling (still small sample size): Add credibility when potential purposeful sample is larger than one can handle; reduces bias within a purposeful category (not for generalizations or representativeness).
- Sampling politically important cases: Attract attention to the evaluation (or avoid attracting undesired attention by purposefully eliminating politically sensitive cases from the sample).
- Combination or mixed purposeful sampling: Triangulation; flexibility; meet multiple interests and needs.

☐ *Determine sample size:* No formula exists to determine sample size. There are trade-offs between depth and breadth, between doing fewer cases in greater depth, or more cases in less depth, given limitations of time and money. Whatever the strategy, a rationale will be needed. Options include:

- Sample to the point of redundancy (not learning anything new).
- Emergent sampling design; start out and add to the sample as fieldwork progresses.
- Determine the sample size and scope in advance.

☐ *Data collection methods:* Determine the mix of observational fieldwork, interviewing, and document analysis to be done in the evaluation. This is not done rigidly, but rather as a way to estimate allocation of time and effort and to anticipate what data will be available to answer key questions.

☐ *Resources available:* Determine the resources available to support the inquiry, including:

- financial resources
- time
- people resources
- access, connections

5. Where fieldwork is part of the evaluation, determine how to approach the fieldwork.

The purpose of field observations is to take the reader into the setting (e.g., program) that was observed. This means that observational data must have depth and detail. The data must be descriptive—sufficiently descriptive that the reader can understand what occurred and how it occurred. The observer's notes become the eyes, ears, and perceptual senses for the reader. The descriptions must be factual, accurate, and thorough without being cluttered by irrelevant minutiae and trivia. The basic criterion to apply to a recorded observation is the extent to which the observation permits the primary intended users to enter vicariously into the program being evaluated.

Likewise, interviewing skills are essential for the observer because, during fieldwork, you will need and want to talk with people, whether formally or informally. Participant observers gather a great deal of

information through informal, naturally occurring conversations. Understanding that interviewing and observation are mutually reinforcing qualitative techniques is a bridge to understanding the fundamentally people-oriented nature of qualitative inquiry.

- ☐ Design the fieldwork to be clear about the role of the observer (degree of participation); the tension between insider (emic) and outsider (etic) perspectives; degree and nature of collaboration with co-researchers; disclosure and explanation of the observer's role to others; duration of observations (short versus long); and focus of observation (narrow vs. broad).

• Role of the Evaluation Observer:	Full participant in the setting	_____/ Part Participant/Part Observer	Onlooker observer (spectator)
• Insider Versus Outsider Perspective:	Insider (emic) perspective dominant	_____/ Balance	Outsider (etic) perspective dominant
• Who Conducts the Inquiry:	Solo evaluator, teams of professionals	_____/ Variations in Collaboration and Participatory Research	People being studied
• Duration of Observations and Fieldwork:	Short, single observation (e.g., 1 site, 1 hour)	_____/ Ongoing Over Time	Long-term, multiple observations (e.g., months, years)
• Focus of Observations:	Narrow focus: single element	_____/ Evolving, Emergent	Broad focus: holistic view
• Use of Predetermined Sensitizing Concepts	Heavy use of guiding concepts to focus fieldwork	_____/ Combination of Focus and Openness	Open: Little use of guiding concepts

- ☐ Be descriptive in taking field notes. Strive for thick, deep, and rich description.
- ☐ Stay open. Gather a variety of information from different perspectives. Be opportunistic in following leads and sampling purposefully to deepen understanding. Allow the design to emerge flexibly as new understandings open up new paths of inquiry.
- ☐ Cross-validate and triangulate by gathering different kinds of data: observations, interviews, documents, artifacts, recordings, and photographs. Use multiple and mixed methods.

- ☐ Use quotations; represent people in their own terms. Capture participants' views of their experiences in their own words.
- ☐ Select key informants wisely and use them carefully. Draw on the wisdom of their informed perspectives, but keep in mind that their perspectives are selective.
- ☐ Be aware of and strategic about the different stages of fieldwork.
    - Build trust and rapport at the entry stage. Remember that the observer is also being observed and evaluated.
    - Attend to relationships throughout fieldwork and the ways in which relationships change over the course of fieldwork, including relationships with hosts, sponsors within the setting, and co-researchers in collaborative and participatory research.
    - Stay alert and disciplined during the more routine, middle phase of fieldwork.
    - Focus on pulling together a useful synthesis as fieldwork draws to a close. Move from generating possibilities to verifying emergent patterns and confirming themes.
    - Be disciplined and conscientious in taking detailed field notes at all stages of fieldwork.
    - Provide formative feedback as part of the verification process of fieldwork. Time that feedback carefully. Observe its impact.
- ☐ Be as involved as possible in experiencing the program setting as fully as is appropriate and manageable while maintaining an analytical perspective grounded in the purpose of the evaluation.
- ☐ Separate description from interpretation and judgment.
- ☐ Be reflective and reflexive. Include in your field notes and reports your own experiences, thoughts, and feelings. Consider and report how your observations may have affected the observed as well as how you may have been affected by what and how you've participated and observed. Ponder and report the origins and implications of your own perspective.

6. Where open-ended interviewing is part of the evaluation, determine how to approach the interviews.

In-depth, open-ended interviewing is aimed at capturing interviewees' experiences with and perspectives on the program being evaluated to facilitate interview participants expressing their program experiences and judgments *in their own terms*. Since a major part of what is happening in a program is provided by people in their own terms, the evaluator must find out about those terms rather than impose upon them a preconceived or outsider's scheme of what they are about. It is the interviewer's task to find out what is fundamental or central to the people being interviewed, to capture their stories and their worldviews.

- ☐ *Types of interviews:* Distinguish and understand the differences between structured, open-ended interviews; interview guide approaches; conversational interviews; and group interviews, including focus groups:

- Structured, open-ended interviews—standardized questions to provide each interviewee with the same stimulus and to coordinate interviewing among team members.
- Interview guide approaches—identifies topics, but not actual wording of questions, thereby offering flexibility.
- Conversational interviews—highly interactive; interviewer reacts as well as shares to create a sense of conversation.
- Focus groups—interviewer becomes a facilitator among interviewees in a group setting where they hear and react to one another's responses.

☐ *Types of questions:* Distinguish and understand the different types of interview questions and sequence the interview to get at the issues that are most important to the evaluation's focus and intended users.

☐ *Listen carefully to responses*: Interviewing involves both asking questions and listening attentively to responses. Using the matrix below if you ask an experiential question, listen to be sure you get an experiential response.

A Matrix of Question Options

	Past	Present	Future
Behaviors/Experiences			
Opinions/Values			
Feelings/Emotions			
Knowledge			
Sensory			
Background			

☐ When *tape-recording*, monitor equipment to make sure it is working properly and not interfering with the quality of responses.

☐ *Practice interviewing* to develop skill. Get feedback on technique.

☐ *Adapt interview techniques* to the interviewee, e.g., children, key informants, elderly who may have trouble hearing, people with little education, those with and without power, those with different stakes in the evaluation findings.

☐ *Observe the interviewee:* Every interview is also an observation.

☐ *Use probes* to solicit deeper, richer responses. Help the interviewee understand the degree of depth and detail desired through probing and reinforcement for in-depth responses.

☐ *Honor the interviewee's experience and perspective.* Be empathetic, neutral, nonjudgmental, and appreciative.

7. Design the evaluation with careful attention to ethical issues.

Qualitative studies pose some unique ethical challenges because of the often emergent and open-ended nature of the inquiry and because of the direct personal contact between the evaluator and people observed or interviewed.

- ☐ *Explaining purpose*: How will you explain the purpose of the evaluation and methods to be used in ways that are accurate and understandable?
  - What language will make sense to participants in the study?
  - What details are critical to share? What can be left out?
  - What's the expected value of your work to society and to the greater good?
- ☐ *Promises and reciprocity:* What's in it for the interviewee?
  - Why should the interviewee participate in the interview?
  - Don't make promises lightly, e.g., promising a copy of the tape recording or the report. *If you make promises, keep them.*
- ☐ *Risk assessment:* In what ways, if any, will conducting the interview put people at risk? How will you describe these potential risks to interviewees? How will you handle them if they arise?
  - psychological stress
  - legal liabilities
  - in evaluation studies, continued program participation (if certain things become known)
  - ostracism by peers, program staff, or others for talking
  - political repercussions
- ☐ *Confidentiality:* What are reasonable promises of confidentiality that can be fully honored? Know the difference between confidentiality and anonymity. (Confidentiality means you know, but won't tell. Anonymity means you don't know, as in a survey returned anonymously.)
  - What things can you *not* promise confidentiality about, e.g., illegal activities, evidence of child abuse or neglect?
  - Will names, locations, and other details be changed? Or do participants have the option of being identified? (See discussion of this in the text.)
  - Where will data be stored?
  - How long will data be maintained?
- ☐ *Informed consent:* What kind of informed consent, if any, is necessary for mutual protection?
  - What are your local Institutional Review Board (IRB) guidelines and requirements or those of an equivalent committee for protecting human subjects in research?
  - What has to be submitted, under what time lines, for IRB approval, if applicable?
- ☐ *Data access and ownership:* Who will have access to the data? For what purposes?

- Who owns the data in an evaluation? (Be clear about this in the contract.)
- Who has right of review before publication? For example, of case studies, by the person or organization depicted in the case; of the whole report, by a funding or sponsoring organization?

☐ *Interviewer mental health*: How will you and other interviewers likely be affected by conducting the interviews?

- What might be heard, seen, or learned that may merit debriefing and processing?
- Who can you talk with about what you experience without breeching confidentiality?
- How will you take care of yourself?

☐ *Advice:* Who will be the researcher's confidant and counselor on matters of ethics during a study? (Not all issues can be anticipated in advance. Knowing who you will go to in the event of difficulties can save precious time in a crisis and bring much-needed comfort.)

☐ *Data collection boundaries:* How hard will you push for data?

- What lengths will you go to in trying to gain access to data you want? What won't you do?
- How hard will you push interviewees to respond to questions about which they show some discomfort?

☐ *Ethical versus legal:* What ethical framework and philosophy informs your work and assures respect and sensitivity for those you study beyond whatever may be required by law?

- What disciplinary or professional code of ethical conduct will guide you?
- Know the Joint Committee Standards (see the Program Evaluation Standards at http://www.jcsee.org/program-evaluation-standards-statements, especially the Propriety standards).

8. Anticipate analysis—design the evaluation data collection to facilitate analysis.

☐ *Design the evaluation to meet deadlines.* Qualitative analysis is labor intensive and time-consuming. Leave sufficient time to do rigorous analysis. Where collaborative or participatory approaches have been used, provide time for genuine collaboration in the analysis.

☐ *Stay focused on the primary evaluation questions and issues.* The open-ended nature of qualitative inquiry provides lots of opportunities to get sidetracked. While it is important to explore unanticipated outcomes, side effects, and unexpected consequences, do so in relation to primary issues related to program processes, outcomes, and impacts.

☐ *Know what criteria will be used by primary intended users to judge the quality of the findings.*

- Traditional research criteria, e.g., rigor, validity, reliability, generalizability, triangulation of data types and sources
- Evaluation standards: utility, feasibility, propriety, accuracy

- Nontraditional criteria: trustworthiness, diversity of perspectives, clarity of voice, credibility of the inquirer to primary users of the findings

☐ *Be prepared for the creativity, ambiguities, and challenges of analysis.* Qualitative inquiry generates a great volume of data—lengthy descriptions of observations and detailed transcripts from interviews.

☐ *Protect the data.* Being out in the field, in the world where programs are taking place, provides lots of opportunities to misplace or lose data. This can threaten promises of confidentiality as well as undermine the credibility of the evaluation. Fieldwork requires being well organized to label the voluminous data obtained (so as to know when, where, and from whom it was gathered), keep it organized, and maintain it securely.

9. Analyze the data so that the qualitative findings are clear, credible, and address the relevant and priority evaluation questions and issues.

☐ Purpose guides analysis. Keep the analysis focused on primary evaluation questions.

☐ Be sensitive to stages and sequence of analysis.

- *Generative and emergent stage*: Ideas for making sense of the data that emerge while still in the field constitute the beginning of analysis; they are part of the record of field notes. In the course of fieldwork, ideas about directions for analysis will occur. Patterns take shape. Possible themes spring to mind. Hypotheses emerge that inform subsequent fieldwork. Record these.
- *Confirmatory Stage:* Later stages of fieldwork bring closure by moving toward confirmatory data collection—deepening insights into and confirming (or disconfirming) patterns that seem to have *appeared.*
- *Systematic analysis following fieldwork:* Write case studies and conduct cross-case analyses based on rigorous review of field notes, interview transcripts, and document analysis.

☐ Purpose guides reporting.

- *Summative evaluations* will be judged by the extent to which they contribute to making decisions about a program or intervention, usually decisions about overall effectiveness, continuation, expansion, and/or replication at other sites. A full report presenting data, interpretations and recommendations typically is required.
- *Formative evaluations* conducted for program improvement may or may not require a detailed, written report for dissemination. Findings may be reported primarily orally. Summary observations may be listed in outline form or an executive summary may be written, but the time lines for formative feedback and the high costs of formal report writing may make a full, written report unnecessary. Staff and funders often want the insights of an outsider who can interview program participants effectively, observe what goes on in the program, and provide helpful feedback. The methods are qualitative, the purpose is practical, and the analysis is done throughout fieldwork; no written report is expected beyond a final outline of observations and implications.

☐ Organize the voluminous data from fieldwork. Inventory what you have.

- Are the field notes complete?
- Are there any parts that you put off to write later and never got to that need to be finished, even at this late date, before beginning analysis?
- Are there any glaring holes in the data that can still be filled by collecting additional data before the analysis begins?
- Are all the data properly labeled with a notation system that will make retrieval manageable? (Dates, places, interviewee identifying information, etc.)
- Are interview transcriptions complete?
- Check out the quality of the information collected.
- Back up the data.

☐ Determine which, if any, computer-assisted qualitative data management and analysis will be used. (Qualitative software programs facilitate data storage, coding, retrieval, comparing, and linking—but human beings do the analysis.) Checklist of considerations in selecting software:

- How you enter your data (typing directly, imported from word processing, scanning; flexible or fixed formatting)
- Storage differences (internal versus external databases)
- Coding variations (on-screen coding versus assigning the codes first)
- Differences in ease of organizing, reorganizing and relabeling codes
- Variations in whether memos and annotations can be attached to codes (especially important for team analysis)
- Data-linking mechanisms and ease vary (connecting different data sources or segments during analysis)
- Ease of navigating and browsing
- Ease, speed, and process of search and retrieval
- Important display variations (e.g., with and without context)
- Tracking details (recording what you've done for review)

☐ Distinguish description, interpretation, and judgment.

- Aim for "thick" description—sufficient detail to take the reader into the setting being described. Description forms the bedrock of all qualitative reporting.
- Use direct quotations so that respondents are presented in their own terms and ways of expressing themselves.
- Keep quotations and field incident descriptions in context.
- Assure that interpretations follow from the qualitative data. Qualitative interpretation begins with elucidating meanings. The analyst examines a story, a case study, a set of interviews, or a

collection of field notes and asks: What does this mean? What insights and answers are provided about central evaluation questions?

- Make the basis for judgments explicit.

☐ Distinguish and separate case studies from cross-case analysis.

- Make cases complete. A case study consists of all the information one has about each case: interview data, observations, the documentary data (e.g., program records or files, newspaper clippings), impressions and statements of others about the case, and contextual information—in effect, all the information one has accumulated about each particular case goes into that case study.
- Make case studies *holistic* and *context sensitive*. Case analysis involves organizing the data by specific cases (individuals, groups, sites, communities, etc.) for in-depth study and comparison. The qualitative analyst's first and foremost responsibility consists of doing justice to each individual case. (Each case study in a report stands alone, allowing the reader to understand the case as a unique, holistic entity. At a later point in analysis it is possible to compare and contrast cases, but initially each case must be represented and understood as an idiosyncratic manifestation of the evaluation phenomenon of interest.)
- Ground cross-case analysis in the individual case studies.
- Identify cross-case patterns and themes with citations and illustrations from the case studies.

☐ Distinguish inductive from deductive qualitative analysis.

- *Inductive analysis* involves *discovering* patterns, themes, and categories in one's data. Findings emerge out of the data through the analyst's interactions with the data.
- *Deductive analysis* involves analyzing data according to an existing framework, e.g., the program's logic model.
- *Build on the strengths of both kinds of analysis.* For example, once patterns, themes, and/or categories have been established through inductive analysis, the final, confirmatory stage of qualitative analysis may be deductive in testing and affirming the authenticity and appropriateness of the inductive content analysis, including carefully examining deviant cases or data that don't fit the categories developed.

☐ Distinguish convergence and divergence in coding and classifying.

- In developing codes and categories, begin with convergence—figuring out what things fit together. Begin by looking for *recurring regularities* in the data. These regularities reveal patterns that can be sorted into categories.
  - Judge categories by two criteria: internal homogeneity and external heterogeneity. The first criterion concerns the extent to which the data that belong in a certain category cohere in a meaningful way. The second criterion concerns the extent to which differences among categories are clear.

    - Prioritize categories by utility, salience, credibility, uniqueness, heuristic value, and feasibility.
    - Test the category system or set of categories for completeness and coherence. The individual categories should be consistent; a set of categories should comprise a whole picture. The set should be reasonably inclusive of the qualitative data and information collected and analyzed.
    - Test the credibility and understandability of the categories with someone not involved in the analysis. Do the categories make sense?
- After analyzing for convergence, the mirror analytical strategy involves examining *divergence.* This is done by processes of extension (building on items of information already known), bridging (making connections among different items), and surfacing (proposing new information that ought to fit and then verifying its existence). The analyst brings closure to the process when sources of information have been exhausted, when sets of categories have been saturated so that new sources lead to redundancy, and when clear regularities have emerged that feel integrated.
    - Carefully and thoughtfully consider data that do not seem to fit including *deviant cases* that don't fit the dominant identified patterns.
- Integrate the analysis. This sequence, convergence then divergence, should not be followed mechanically, linearly, or rigidly. The processes of qualitative analysis involve both technical and creative dimensions.

☐ Construct a process-outcomes matrix for the program.

- Distinguish process descriptions from outcome documentation.
- Show linkages between processes and outcomes.

☐ Integrate and reconcile qualitative and quantitative findings as appropriate.

- Note where qualitative and quantitative findings reinforce one another.
- Note and explain differences.

☐ Use strategies to enhance the rigor and credibility of analysis.

- Consider and discuss alternative interpretations of the findings.
- Carefully consider and discuss cases and data that don't fit overall patterns and themes.
- Triangulate the analysis. Options include:
    - Check out the consistency of findings generated by different data-collection methods, i.e., *methods triangulation.*
    - Check out the consistency of different data sources within the same method, i.e., *triangulation of sources.*
    - Use multiple analysts to review findings, i.e., *analyst triangulation.*

- Use multiple perspectives or theories to interpret the data, i.e., theory/perspective triangulation.

☐ Determine substantive significance. In lieu of statistical significance, qualitative findings are judged by their *substantive significance*. The analyst makes an argument for substantive significance in presenting findings and conclusions, but readers and users of the analysis will make their own value judgments about significance. In determining substantive significance, the analyst addresses these kinds of questions:

- How solid, coherent, and consistent is the evidence in support of the findings? (Triangulation, for example, can be used in determining the strength of evidence in support of a finding.)
- To what extent and in what ways do the findings increase and deepen understanding of the program being evaluated?
- To what extent are the findings consistent with other knowledge? (A finding supported by and supportive of other work has confirmatory significance. A finding that breaks new ground has discovery or innovative significance.)
- To what extent are the findings useful for the evaluation's purpose?

☐ Determine if an expert audit or metaevaluation is appropriate because the stakes for the evaluation are high (e.g., major summative evaluation) and the credibility of the qualitative findings will be enhanced by external review.

- Quality audit: An external audit by a disinterested expert can render judgment about the quality of data collection and analysis.
  - An audit of the qualitative data collection process results in a *dependability judgment.*
  - An audit of the analysis provides a *confirmability judgment.*
  - Any audit would need to be conducted according to appropriate criteria given the evaluation's purpose and intended uses.
- A metaevaluation employs general evaluation standards for the review. (See the Program Evaluations Standards at http://www.jcsee.org/program-evaluation-standards-statements

10. Focus the qualitative evaluation report.

☐ Determine what is *essential* and make that the focus of the evaluation report and other communication strategies (e.g., oral reports).

- Focus on what will be most useful and meaningful. Even a comprehensive report will have to omit a great deal of information collected by the qualitative evaluator. Evaluators who try to include everything risk losing their readers in the sheer volume of the presentation. To enhance a report's impact, the evaluation should address each major evaluation question clearly, that is, succinctly present the descriptive findings, analysis, and interpretation of each focused issue.
- An evaluation report should be readable, understandable, and relatively free of academic jargon.

☐ Review the evaluation findings to distinguish and serve three functions:

- Confirm and highlight major evaluation findings supported by the qualitative data.
- Disabuse people of misconceptions about the program.
- Illuminate important things not previously known or understood that should be known or understood.

☐ Determine the criteria by which the evaluation report will be judged based on intended users and evaluation purposes and meet those criteria to enhance the report's credibility and utility. Alternative criteria frameworks include:

- Traditional scientific research criteria
- Constructivist criteria
- Artistic and evocative criteria
- Critical change criteria
- Pragmatic criteria
- Evaluation standards (See standards checklist)

References

Greene, J. C. (2000). Understanding social programs through evaluation. In N. K. Denzin & Y. S. Lincoln (Eds.). *Handbook of qualitative research* (2nd ed.) (pp. 981- 999). Thousand Oaks, CA: Sage.

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed). Thousand Oaks, CA: Sage.

Rossman, G. B., & Rallis, S. F. (1998). *Learning in the field: An introduction to qualitative research.* Thousand Oaks, CA: Sage.

Schwandt, T. A. (2001). *Dictionary of qualitative inquiry* (2nd revised ed.) Thousand Oaks, CA: Sage.

Stake, R. (1995). *The art of case study research*. Thousand Oaks, CA: Sage.

Williams, D. D. (Ed.) (1986). Naturalistic evaluation. *New Directions for Program Evaluation, 30.*

Suggested Citation

Patton, M. Q. (2003). *Qualitative evaluation checklist*. Retrieved from http://wmich.edu/evaluation/checklists

This checklist is provided as a free service to the user. The provider of the checklist has not modified or adapted the checklist to fit the specific needs of the user and the user must use their own discretion and judgment in using the checklist. The provider of the checklist makes no representations or warranties that this checklist is fit for the particular purpose contemplated by the user and specifically disclaims any such warranties or representations.