Methods of Rapid Evaluation, Assessment, and Appraisal

2 authors, including:

Pennie G Foster-Fishman
Michigan State University
52 PUBLICATIONS   **2,842** CITATIONS

American Journal of Evaluation

Methods of Rapid Evaluation, Assessment, and Appraisal

Miles McNall and Pennie G. Foster-Fishman

The online version of this article can be found at:

Methods of Rapid Evaluation, Assessment, and Appraisal

Miles McNall
Pennie G. Foster-Fishman
Michigan State University

Abstract: A central issue in the use of rapid evaluation and assessment methods (REAM) is achieving a balance between speed and trustworthiness. In this article, the authors review the key differences and common features of this family of methods and present a case example that illustrates how evaluators can use rapid evaluation techniques in their own work. In doing so, the authors hope to (a) introduce readers to a family of techniques with which they may be unfamiliar, (b) highlight their strengths and limitations, and (c) suggest appropriate contexts for use. Ultimately, the authors hope that REAM becomes a valuable addition to evaluators' toolkits.

Keywords: *rapid evaluation; real-time evaluation; rapid assessment*

The evaluation of ongoing responses to humanitarian emergencies presents an extreme example of an urgent need for evaluation findings. In response to the unfolding crisis in Kosovo in 1998 to 1999, Danish International Development Assistance conducted what is believed to be the first real-time evaluation (RTE) of a humanitarian emergency response. Since then, the Evaluation and Policy Analysis Unit of the United Nations High Commissioner for Refugees (UNHCR) has used RTE to assess responses to humanitarian crises in Sudan-Eritrea, Angola, Pakistan, Afghanistan-Iran, and Chad (Bartsch & Belgacem, 2004; Sandison, 2003). Although most evaluations do not have life and death consequences, there are occasions where the timely provision of information can, if not forestall catastrophe, at least mitigate its impact. The dramatically improved preparation for and response to Hurricane Rita as compared to Hurricane Katrina demonstrates that a quick postmortem is often sufficient to vastly improve the effectiveness of subsequent responses. On other–less dire–occasions, communities or organizations need quick data collection and analyses to assist in urgent decision making, grant applications, and policy-making processes.

Over the last two decades, several methods of rapid evaluation, assessment, and appraisal have emerged, including, but not limited to, RTE (Bartsch & Belgacem, 2004; Jamal & Crisp, 2002; Sandison, 2003), rapid-feedback evaluation (RFE; McNall, Welch, Ruh, Mildner, & Soto, 2004; Sonnichsen, 2000; Wholey, 1983), rapid assessment (RA; Beebe, 2001; Trotter & Singer, 2005), rapid ethnographic assessment (REA; Bentley et al., 1988; Guerrero et al., 1999; Kresno, Harrison, Sutrisna, & Reingold, 1994), rapid evaluation methods (REM; Anker, Guidotti, Orzeszyna, Sapirie, & Thuriax, 1993), rapid rural appraisal (RRA; Chambers, 1994a; Rifkin, 1992), and participatory rural appraisal (PRA; Chambers, 1994a). Although the origins,

methods, and contexts of practice for each of these approaches are distinct, what they share in common is a similar set of techniques for putting trustworthy, actionable information in the hands of decision makers at critical moments.

A central issue in the use of rapid evaluation and assessment methods (REAM) is achieving a balance between speed and trustworthiness. In this article, we review the common features and major differences among this family of methods. Then, we present a case example that illustrates how evaluators can use rapid evaluation techniques in their work to produce trustworthy findings in a matter of months, if not weeks. In doing so, we hope to (a) introduce readers to a family of techniques with which they may be unfamiliar, (b) highlight their strengths and limitations, and (c) suggest appropriate contexts for use. Ultimately, we hope that REAM becomes a valuable addition to evaluators' toolkits. In the following section, we briefly review the major lineages of REAM before discussing the core principles and methods they share. As we review the lineages of REAM, we will not engage in a debate over the meaning of the terms *evaluation*, *assessment*, and *appraisal*. Instead, we will describe the lineages using their own labels.

Lineages

In addition to several articles and books describing various approaches to rapid evaluation, assessment, or appraisal, our review of the literature found descriptions of 13 projects published between 1988 and 2005 that used these methods in a field setting. A summary of representative examples of each major type of REAM, including each project's purpose, target population, methods, and extent of stakeholder participation in the evaluation is presented in Table 1. We do not claim that this summary is exhaustive of all such published projects but do feel that it provides a fair representation of the major types of approaches to rapid evaluation, assessment, or appraisal employed in field settings in the last two decades.

Real-Time Evaluations

Real-time evaluations developed in the 1990s in response to the proliferation of humanitarian crises. The UNHCR needed a method for rapidly evaluating the effectiveness and impact of its operational responses to emergencies and ensuring that findings were "used as an immediate catalyst for organizational and operational change" (UNHCR, 2002, p. 43).

Real-time evaluations are initiated as soon as a new crisis emerges or appears imminent. Initially, evaluators are part of the crisis response "cell," systematically collecting and reviewing data as the crisis unfolds (Jamal & Crisp, 2002). As part of this cell, RTE evaluators interact and share their observations and recommendations on an ongoing basis with field staff so as to allow operational problems to be quickly corrected and potential problems to be avoided. Typically, RTEs use a mixed-methods approach, involving semistructured interviews, site visits, a limited number of in-depth interviews, focus groups, and reviews of secondary documentation (Sandison, 2003). At the conclusion of an RTE, evaluators hold an interactive debriefing session with UNHCR field staff and representatives of partner organizations (UNHCR, 2002).

An evaluation of the UNHCR's response to the humanitarian crisis developing in eastern Chad in 2003 illustrates the sorts of questions addressed by an RTE. This RTE assessed the extent to which the UNHCR's humanitarian response had the following features: (a) refugees had access to life-sustaining assistance, (b) UNHCR's response mode had shifted from reactive to proactive planning, (c) effective refugee protections were in place, (d) an effective management structure had been established, and (e) viable implementation arrangements had been made (Bartsch & Belgacem, 2004).

Table 1
Key Features of Rapid Evaluation and Assessment Methods Projects

Approach	Authors	Purpose	Target Population	Methods	Stakeholder Participation in Evaluation
Participatory rural appraisal	Rutta et al. (2005)	To assess refugee perceptions of the quality of health care in order to improve the quality of that health care	Residents of refugee camps in Ngara, Tanzania	Focus groups Free listing Problem ranking Historical narration Transect walks	None mentioned
Rapid assessment	Aral, St. Lawrence, Dyatlov, & Kozlov (2005)	To assess the relative contributions of drug use and sex work to the transmission of sexually transmitted infections in St. Petersburg, Russia in order to develop effective prevention interventions.	Sex workers and drug users in St. Petersburg, Russia	Interviews Observation Mapping Analysis of existing data	None mentioned
	Desmond et al. (2005)	To develop a typology of groups at risk of HIV/STI based on local cultural understandings to inform the development and implementation of an HIV prevention program	Female recreation facility workers and male mineworkers in a mining town in northwestern Tanzania	Interviews Observation	None mentioned
	Kirsch (1995)	To generate information on population groups at risk in order to design communication and education programs and develop prevention materials	Adolescents 13 to 18 years of age in four cities in Brazil, Mexico, and Paraguay	Focus groups Surveys	Community leaders participated in defining the problem and generating solutions
Rapid ethnographic assessment	Bentley et al. (1988)	To provide information on peoples' beliefs about diarrhea in order to design effective interventions to in order control it	Residents of rural Andean villages and one urban center in Peru Unspecified location in Nigeria	Interviews	None mentioned

(continued)

153

Table 1 (continued)

Approach	Authors	Purpose	Target Population	Methods	Stakeholder Participation in Evaluation
	Guerrero et al. (1999)	To determine factors associated with the absence of exclusive breastfeeding in order to design effective campaigns to promote exclusive breastfeeding	Mothers of children under 5 years of age in periurban Mexico City	Interviews Questionnaires	None mentioned
	Kresno, Harrison, Sutrisna, & Reingold (1994)	To identify local beliefs, perceptions, and practices surrounding acute respiratory infections in infants and young children	Mothers with children 5 years of age	Interviews	None mentioned
Rapid evaluation methods	Anker, Guidotti, Orzeszyna, Sapirie, & Thuriax (1993)	To assess the quality of health care services, identify operational problems, and assist managerial action	Patients	Interviews Focus groups Observation Record reviews	Health staff participated in identification of issues and preparation of data collection instruments
Rapid-feedback evaluation	McNall et al. (2004)	To identify reasons for poor recruitment and retention rates in an integrated HIV/AIDS care intervention	Ambulatory care patients with HIV/AIDs and co-occurring substance abuse and mental health disorders	Interviews Analysis of existing data	None mentioned
Real-time evaluation	Bartsch & Belgacem (2004)	To evaluate the effectiveness of UNHCR's response to a humanitarian crisis in Chad	Sudanese refugees in Chad	Interviews	None mentioned

Note: UNHCR = United Nations High Commissioner for Refugees.

a. By target population, we mean the population of individuals who ultimately stand to benefit from the application of the knowledge gained from rapid evaluation and assessment methods projects.

RFE

The idea for RFE grew out of Joseph Wholey's experiences as an evaluator working with federal program managers and his awareness of clients' needs for the timely delivery of findings to inform key programming decisions. In the view of Wholey and his colleagues, rough approximations delivered at the right time are better than precise results delivered too late for decision makers to act on them (Bellavita, Wholey, & Abramson, 1986). An RFE involves the use of existing program data to make a quick, preliminary assessment of program performance (Sonnichsen, 2000; Wholey, 1983). Wholey's (1983) RFE model consists of five basic steps:

1. Collection of existing data on program performance
2. Collection of new data on program performance (typically limited to brief interviews with program staff)
3. Preliminary evaluation
4. Development and analysis of alternative designs for full-scale evaluation
5. Assisting policy and management decisions

Although Step 4 suggests that an RFE is simply a prelude to a full-scale evaluation, Sonnichsen (2000) argues that, in some cases, Step 4 can be skipped: "In many instances, the information gathered during rapid-feedback evaluations may be sufficient to answer client questions and no additional evaluation will be necessary" (p. 218).

Others have pursued the idea of RFE as a stand-alone approach for providing quick answers to highly focused questions (McNall et al., 2004; Sonnichsen, 2000). As a stand-alone approach, RFEs are likely to be employed in situations where program managers have either tightly focused questions about program performance or have identified a problem with program operations and need more information to decide how to correct the problem.

RA

The techniques of RA are derived primarily from the traditions of ethnography, action research (Argyris, Putnam, & Smith, 1985; Lewin, 1948), and participatory action research (Greenwood, Whyte, & Harkavy, 1993). In comparison to traditional ethnographic methods, where extended immersion in a cultural context is the norm, RA is more rapid, cost effective, technically eclectic, and pragmatic (Vincent, Allsop, & Shoobridge, 2000). In RA, teams of researchers are deployed to gather information from small samples of key informants and local residents using surveys, semistructured interviews, focus groups, transect walks, and mapping (Garrett & Downen, 2002). Existing data sets are also exploited to provide a more comprehensive picture of the problem (Aral, St. Lawrence, Dyatlov, & Kozlov, 2005; Vincent et al., 2000). Like rapid-feedback evaluation, the primary purpose of RA is to quickly generate information to assist decision making; however, in contrast to rapid-feedback evaluation, RA is more often used to generate information about health and social problems to aid in the design of culturally appropriate interventions for health and social problems than to evaluate existing programs (Trotter & Singer, 2005; Vincent et al., 2000).

Since the 1980s, RA has been used in public health initiatives worldwide (Trotter & Singer, 2005; Scrimshaw & Gleason, 1992). Rapid assessment has been used to study sexually transmitted infections (STI) in St. Petersburg, Russia (Aral et al., 2005); HIV and STI transmission in a gold mining town in Tanzania (Desmond et al., 2005); the HIV/AIDS crisis in racial and ethnic minority communities in three U.S. cities (Needle et al., 2003); malaria in Indonesia (Utarini, Winkvist, & Ulfa, 2003); drug use in South Australia (Vincent et al., 2000); tuberculosis control in KaNgwane, South Africa (Lee & Price, 1985); adolescents' risk of drug addiction in four Latin American cities (Kirsch, 1995); and refugees' perceptions of the quality of health

care in Ngara, Tanzania (Rutta et al., 2005). Also, RA has been used to identify the programming needs of the urban poor in Bangladesh and Tanzania (Garrett & Downen, 2002).

Rapid assessment has a highly pragmatic, action-oriented nature. Of the eight RA projects published between 1995 and 2005 reviewed for this article, all gathered information about a social or health issue and used that information either to improve an existing program or service or to design more effective and culturally appropriate interventions. Rapid assessment projects were less consistent in the degree to which either members of targeted beneficiary populations or representatives of local organizations were involved in the design or implementation of RA projects. Of the eight published descriptions of RA projects, four make no mention of participatory practices, although the rest describe forms of participation ranging from soliciting the input of targeted beneficiary groups and local experts to involving community members in the design and execution of projects and the reporting and interpretation of results.

REA

Like RA, REA was developed to provide quick assessments of local conditions to inform the design of effective interventions. Also similar to RA, the major focus of REA is on public health in developing nations. Rapid ethnographic assessment has been used to develop effective diarrhea control programs in Peru and Nigeria (Bentley et al., 1988); to determine the factors associated with the absence of exclusive breastfeeding in periurban Mexico City (Guerrero et al., 1999); and to understand local beliefs, perceptions, and practices regarding acute respiratory infections in West Java, Indonesia (Kresno et al., 1994).

The major differences between RA and REA is that REA tends to make use of a more limited range of research methods and to be more exclusively focused on exploring indigenous understandings of health issues than does RA. Among the three published REA projects mentioned above, research methods were limited to key informant interviews, questionnaires, and use of existing data sets. This is not to say that a greater understanding of local beliefs and attitudes regarding health issues is not valuable in its own right; such understandings are, in fact, critical to developing effective interventions (Bentley et al., 1988).

Rapid Evaluation Methods

Rapid evaluation methods were developed by the World Health Organization for the purposes of assessing the quality of health care services, identifying operational problems, and assisting managers in taking corrective action (Anker et al., 1993). Rapid Evaluation Method was born of "a need for a quick, accurate, and economical method of evaluation of facilities and client satisfaction" (Anker et al., 1993, p. 15) and concerns about the practicality and cost effectiveness of surveys.

Like the preceding methods of rapid evaluation and assessment, REM collects data from a variety of sources using several different methods. Data are collected by means of clinic exit interviews with patients; health staff interviews; observations of staff task performance; community and staff focus group discussions; review of clinic records; checking of facilities, equipment, and supplies; and household interviews.

Despite the number of different data collection methods used, REM research teams are still able to rapidly collect, analyze, and report data. For example, in a test of REM in five developing countries (Anker et al., 1993), teams of four to five surveyors were able to complete their evaluations of local health systems in 6 to 10 days. During this time each team was able to visit one provincial health office, one district hospital, two to four health centers, and four to eight villages.

PRA

Employed principally in developing countries in the areas of natural resources management, agriculture, poverty and social programs, and health and food security, PRA is "a family of approaches and methods to enable rural people to share, enhance, and analyze their knowledge of life and conditions, to plan and to act" (Chambers, 1994a, p. 953).

The ancestry of PRA can be traced to five distinct traditions: activist participatory research, agroecosystem analysis, applied anthropology, field research on farming systems, and RRA Chambers, 1994a). The most direct source of PRA is RRA, which evolved in the late 1970s and early 1980s as a corrective to the biased and inaccurate perceptions resulting from what has been called "rural development tourism." Rural development tourism involves brief visits by urban professionals to rural areas, chosen to serve as Potemkin villages[1] (Chambers, 1994a).

In contrast to RA, which is primarily concerned with the gathering of health and disease information as a product, in rural appraisal, the emphasis is on information gathering as a process with defined characteristics, including (Rifkin, 1992):

- Community involvement in the gathering and analysis of data
- A holistic and systematic approach
- Multidisciplinary and interactive methods
- Flexible responses
- An emphasis on communication and listening skills
- Visual display of information

Of the kinds of REAM reviewed so far, PRA has the greatest variety of tools in its toolkit (Chambers, 1994a)—most likely a consequence of the diversity of its lineages. PRA involves the collection of data from a variety of sources, including:

- Secondary sources (files, reports, etc.)
- Key informants
- Local residents
- Observations

PRA also uses a variety of methods for collecting data. A few of the 29 methods that Chambers (1994a) lists are described below.

Semistructured interviews use a few general questions to generate an open, two-way conversation between the interviewer and the respondent.

Group interviews, also known as focus groups, are structured, small-group interviews of a relatively homogeneous group of people (e.g., mothers of children aged 0-2) focused on a particular topic (e.g., breastfeeding).

Oral histories involve recording people's memories about their own experiences.

Transect walks entail walking through an area, observing, and asking questions to identify natural resources and how they are being used.

Mapping and modeling of local conditions involves local people in the construction of maps of local demographics, health, resources, services, or land use.

Timelines and trend change analysis involves local people in the development of chronologies of significant local events and changes in such things as land use, local customs and practices, or population and migration.

Seasonal calendars display variations in local conditions, including rain, crops, labor, diet, illness, etc.

Although PRA employs a variety of methods, not all are used in every appraisal. The particular mix of methods used in any given assessment depends on what is suitable for the particular issue being investigated.

PRA as a participatory process involves both the participation of professional researchers in community life and the participation of local residents in research. In PRA, researchers may perform village tasks while local residents participate in research activities from data collection to the analysis, presentation, and interpretation of results. In contrast to both RA and RRA, where data gathering involves extracting information from community members, data in PRA are "generated, analyzed, owned and shared by local people as part of a process of their own empowerment" (Chambers, 1994b, p. 1253).

Common Methods

Despite their diverse lineages and contexts of application, a comparison of the methods used in the preceding approaches reveals a set of core elements (Table 2).

Methods

All of the approaches reviewed above involved the collection of data from a variety of sources using a diversity of methods. Frequently, both quantitative and qualitative data were collected. Quantitative data were collected primarily through surveys and the review of existing data sets. In the case of the rapid assessment/appraisal studies, existing data sets included epidemiological and sociological data; rapid evaluation approaches involved the use of organizational administrative data sets. Qualitative data were collected through three principal methods: (a) formal and informal interviews with key informants and others, (b) focus groups, and (c) naturalistic observations. Finally, mapping was frequently used to delineate areas and populations affected by the problem in question.

Process

As its name suggests, REAM is indeed rapid, with efforts lasting anywhere from a few days to a few months. In some cases, the work is done in a participatory manner, involving community members in a wide range of project activities. Of the 13 REAM projects we reviewed, the majority (9 projects) did not mention any form of community member participation in the evaluation other than the provision of data. The project that involved the deepest level of community member participation was Rutta et al.'s (2005) PRA of refugee perceptions of the quality of health care. Indeed, participation of the target population is a defining feature of PRA. Beebe's (2001) rapid assessment process also calls for a high level of target population participation in the research and evaluation process. Another characteristic of REAM is that the process is team based, involving the collaboration of members of the team at every step of the process, from planning and data collection to the interpretation of findings and presentation of results. Finally, REAM is performed in an iterative manner, involving the analysis of data while they are still being collected and the use of preliminary findings to guide decisions about additional data collection. This cycle is repeated until new data cease to provide new information of significance to the major research or evaluation questions.

Table 2
Core Elements of Rapid Research and Evaluation Methods

Methods

Mixed methods:
    Quantitative approaches typically include:
        Quantitative surveys
        Review of existing data sets
    Qualitative approaches usually include:
        Key informant interviews
        Focus groups
        Naturalistic observations
        Record reviews
    Mapping of areas affected by problem

Process

Rapid: Evaluation, assessment, or appraisal lasts from a few weeks to a few months

Participatory: Representatives of local populations and institutions are involved in the planning and implementation of the research

Team based: Members of the research team work collaboratively on all aspects of the research process, from planning and data collection to the interpretation of findings and presentation of results.

Iterative: Data are analyzed while they are being collected, and preliminary findings are used to guide decisions about additional data collection. This process continues until theoretical saturation is achieved.

Key Differences

Although the different types of REAM share much in common, a key difference between the assessment/appraisal (RA, REA, and PRA) and evaluation (RTE, RFE, and REM) variants is that the former typically focus on assessing a situation or context within a defined geographic area, whereas the latter tend to focus on projects, programs, or systems. For example, rapid assessment approaches have been used to study HIV transmission patterns in particular cities (Aral et al., 2005) or towns (Desmond et al., 2005), whereas rapid evaluation approaches have been employed to evaluate an HIV/AIDS health care intervention (McNall et al., 2004) and health care facilities in five developing countries (Anker et al., 1993). Because of their focus on geographic areas, rapid assessment/appraisal approaches typically use ethnographic methods to capture the dimensions of local cultures relevant to the particular problem or issue at hand. Because of their focus on projects, programs, and systems, rapid evaluation approaches are primarily concerned with the thoughts and behaviors of the actors within those entities. As a consequence, these approaches tend not to pay as much attention to cultures outside of the projects, programs, or systems of interest. In principle, however, there is no reason why rapid evaluation approaches could not use techniques such as rapid ethnography to assess elements of organizational culture relevant to the problem of interest.

Despite this difference, the contrasts between rapid evaluation and assessment/appraisal techniques should not be overdrawn and are likely artifacts of their distinct intellectual lineages and historical contexts of application. At this point in their developmental history, there has been enough cross-fertilization between the variants of REAM that they have become almost indistinguishable, allowing an identifiable set of core principles and methods to emerge from the confusion of names.

A Case Example of REAM

So how exactly does someone engage in REAM in a manner that stays true to the quick responsive nature of the method while generating results that are believable and useful? In our review of the REAM literature, we found little explicit discussion of methods, processes, and timelines or the criteria used to determine the adequacy of the conclusions drawn. Because evaluators are increasingly being asked to provide "real data in real time," we thought it would be useful to describe the processes and strategies we used in one of our recent REAM cases and highlight how we applied the adequacy criteria proposed by Guba and Lincoln (1989) in their Fourth Generation Evaluation framework to ensure the accuracy and usefulness of our findings. We have found Guba and Lincoln's framework particularly useful for our REAM projects because this approach was designed, in part, for situations where intense collaboration between the evaluator and members of the targeted setting is desired to ensure that evaluation process and findings represent and accurately portray the "claims, concerns, and issues of stakeholders" (p. 50). In the following case study, we illustrate how we have used REAM in an evaluation that lasted 8 weeks from initial design to final reporting and applied Guba and Lincoln's adequacy criteria to ensure the credibility and usefulness of the findings. Table 3 highlights their criteria, summarizes how we applied some of these criteria in our case, and identifies additional strategies we could have employed to apply more fully these criteria in this project. Certainly, not all rapid evaluation users will agree with our selection of Guba and Lincoln's framework. There are other adequacy criteria that could be used, including criteria proposed by Miles and Huberman (1994) and Patton (2002). For example, Patton argues that the credibility of qualitative research depends on the extent to which rigorous methods are used, the credibility of the researcher(s), and a philosophical belief in the value of qualitative inquiry. Our emphasis in this article is on the need for rapid evaluators to select an adequacy criteria framework that is most applicable to their project's purpose and context and to make the process of evaluating their work against these criteria transparent.

Background

The second author was hired by a national training organization to conduct a formative and summative evaluation of its training and technical assistance efforts. Given some upcoming grant application deadlines, only 8 weeks were available to conduct this evaluation (from design to final reporting). The second author hired two additional staff members to assist with the evaluation (both contributing approximately 20-30 hrs per week for the 8-week period). While this team collaborated extensively with the organization's staff in the design and implementation of this work, all data were collected and analyzed by the evaluation team.

Design: Weeks 1 and 2

To ensure that the design of this evaluation fit the local context and the needs of the sponsoring organization and that all parties agreed on the goals of the evaluation, the evaluation team spent 1 day on-site in the local organization. To maximize the utility of this on-site visit, the evaluation team reviewed all relevant secondary documents, including the organization's strategic plan, annual reports, and internal evaluation reports prior to this visit. The day started with a group meeting with all staff where the general purpose of the evaluation was discussed and preliminary thoughts about the evaluation design were considered. Input from staff was requested and modifications to the design and purpose were made accordingly.

Table 3
Description of Guba and Lincoln's Adequacy Criteria and Their
Application to Our REAM Case

Guba & Lincoln's Criteria	Definition	Application in Case Example	Additional Application Strategies
DOMAIN 1: trustworthiness	Quality and rigor of the data and findings. Includes four criteria: credibility, transferability, dependability, and confirmability.		
Trustworthiness criteria: credibility	Extent to which findings accurately portray respondents' constructions. Involves the following: *Prolonged engagement* in targeted site to build rapport and trust between evaluators and setting members and provide evaluators with a deeper understanding of the relevant culture. *Persistent observation* of site to provide sufficient understanding. *Peer debriefing:* Extensive discussions of data and preliminary findings with one or more peers to refine thinking. *Negative case analysis:* The constant reworking of hypotheses in light of disconfirming evidence. *Progressive subjectivity:* Researchers identify and articulate any biases they hold, examine how their understandings shift during the project, and attend to how these biases might affect interpretations. *Member checks* involve sharing and checking findings and interpretations with the people from whom the data were collected.	*Prolonged engagement*: Not viable in REAM. *Persistent observation*: No relevant events available to observe. *Peer debriefing*: Daily, short debriefings between staff; weekly full debriefing sessions at project meetings. *Negative case analysis*: Themes or cases that did not "fit" into emerging understandings were examined. Expanded themes or hypotheses were developed to incorporate these findings. *Progressive subjectivity*: Team members talked frequently about their biases or assumptions. *Member checks*: Questions or inconsistencies were discussed with site staff throughout project. Feedback was sought from site staff on preliminary final report.	*Progressive subjectivity*: Could have recorded initial biases about site following initial visit and revisited these assumptions. *Member checks*: Could have shared and sought feedback on emerging frameworks in subsequent interviews.
Trustworthiness criteria: transferability	Researchers describe features of targeted context in detail and suggest additional contexts to which findings might be generalized.	Extensive background and case information included in final report.	

(continued)

Table 3 (continued)

Guba & Lincoln's Criteria	Definition	Application in Case Example	Additional Application Strategies
Trustworthiness criteria: dependability	Concerned with stability over time in researchers and methods. Assessed by means of a dependability audit, which involves reviewing project records to determine the extent to which project procedures and changes are documented.	Kept a written audit trail of evaluation methods, decisions, and shifts in processes.	
Trustworthiness criteria: confirmability	Extent to which findings are grounded in the data. Assessed by means of confirmability audits, which involve reviewing research records to determine if findings can be traced to data and data to original sources.	Kept all case-summary, substantive theme, and pattern analysis documents. Data and themes in all documents linked to subject IDs.	
DOMAIN 2: Authenticity	Extent to which intent of inquiry is maintained—specifically, claim of accurately representing stakeholders' views.		
Authenticity criteria: fairness	Extent to which different stakeholder perspectives are elicited and taken into account. Involves identifying all stakeholders, soliciting their perspectives, and engaging in open negotiations with them around recommendations and future actions.	Identifying and soliciting stakeholder perspectives: Staff and leaders identified relevant stakeholders. To include a range of perspectives, staff was asked to identify individuals who were satisfied or dissatisfied with the training. Interviewed representatives from all stakeholder groups. *Open negotiation of recommendations:* Not feasible. Instead, all interviewees were asked to provide recommendations, and these were included in the final report.	*Open negotiation:* Could have sent complete list of recommendations to stakeholders via e-mail and asked for their feedback and prioritization.
Authenticity criteria: ontological authenticity	Extent to which stakeholders' perceptions of the world have been improved or expanded.	Staff and leaders reported that the evaluation expanded their understandings of the organization and its programs.	

(continued)

Table 3 (continued)

Guba & Lincoln's Criteria	Definition	Application in Case Example	Additional Application Strategies
Authenticity criteria: educative authenticity	Extent to which individuals have developed a better understanding of other stakeholders' experiences and perspectives.	Differences across stakeholders' experiences with the organization were highlighted in report. Feedback processes suggested that staff learned new stakeholder concerns.	
Authenticity criteria: catalytic authenticity	Extent to which the evaluation elicits action and change.	Year 2 evaluation provided opportunity to track changes linked to Year 1 report.	
Authenticity criteria: tactical authenticity	Extent to which stakeholders feel empowered by the evaluation and by the ability to influence the actions taken.	Not relevant to this project.	

Note: REAM = rapid evaluation and assessment methods.

Following this meeting, one-on-one interviews were held with all staff members (*N*=8). These interviews served to further inform our design (i.e., What do you need to learn from this evaluation to improve your work in this organization?) and to provide background evaluation information (i.e., Tell me about your role in this organization and what you are trying to accomplish.) and formative evaluation information (i.e., What gets in the way of you accomplishing your goals? What could this organization do better?*)*. Midway into these interviews, the evaluation team met to discuss what they were learning and to determine if additional questions needed to be asked of the remaining staff.

Extensive notes were taken during these interviews, and digital recordings were made as backup data sources. To expedite the data analysis process, data recording and analyses were conducted simultaneously. As the team members typed their contact summary sheets for each informant, data were entered into this form as thematic codes along with evidence from the interviews for the themes that were identified. Within 48 hrs of the interviews, all of the data were analyzed, and key themes and issues were identified. Team members met frequently during this 48-hr period to discuss the themes that were emerging to ensure their credibility (peer debriefing).

The information gathered during this stage was used to redesign the evaluation plan and to design the interview protocol for collecting data from various stakeholders. These documents were sent to the executive director for review and approval. To ensure that we adequately represented multiple stakeholder perspectives, we worked closely with site staff to identify all stakeholder groups from which to gather data in this evaluation (fairness).

Data Collection and Analyses: Weeks 2 to 6

Data collection and analyses happened simultaneously during a 4-week period. Phone interviews that lasted between 1 and 2 hrs were conducted with 32 stakeholders. Extensive notes were taken, and digital recordings were done as backup. On the completion of each interview, a case summary sheet was immediately produced in the same manner as described above. Throughout the process, extensive peer debriefing occurred to increase the credibility of the results. The two staff who were conducting the interviews met almost daily to debrief

about the themes they were identifying and issues that were emerging. The whole team met weekly to discuss issues and modify the protocol as needed. A detailed audit trail that explained our evaluation processes and changes (dependability audit) was recorded.

By Week 4, enough interviews had been conducted to begin some cross-case comparisons. Using Miles and Huberman's (1994) matrices approach, cross-case summaries were created for each informant, highlighting themes (and providing supportive evidence) in response to each of the core evaluation questions. On the completion of each interview, data were entered into this matrix and the individual case summary. These matrices became effective tools for helping the team (a) identify patterns across informants, (b) recognize shared themes and larger categories of meaning, and (c) conduct negative case analyses. (See Table 4 for a brief example of the matrices that were used in this project.) Overall, the detailed note taking, case summaries, and matrices, coupled with the linking of all data and themes to subject IDs, served as part of the confirmability audit trail, which demonstrated that our findings were grounded in the data themselves. (See Miles & Huberman, 1994, for a more detailed description of case summaries and matrices.)

Secondary data analysis happened simultaneously, including analyses of client service databases (numbers and types served) and feasibility assessments of the theories of change. The latter included linking current training activities and reported outcomes with the organization's theory of change to determine if (a) the current activities were aligned with the overall goals and mission of the organization and (b) the outcomes that were emerging fit with the organization's model of how training would produce change and what changes would occur.

Report Writing: Weeks 6 to 8

To expedite the report-writing process, we organized it by core evaluation questions. Each team member was assigned a different set of questions to write up. The primary author then reviewed each section, cross-checking these final interpretations against the cross-case matrices and individual case summaries as an additional confirmability audit check. One week before the final report was due, a PowerPoint presentation of the findings was presented to the site staff, who was provided with opportunities to react to the findings. This member checking process allowed the team to determine the accuracy of their conclusions. Key issues or suggestions that emerged were then integrated into the final report.

The content of the report highlighted two core findings: (a) Stakeholders who participated in training sessions found them valuable and transformative, though they offered specific ways the training institute could better meet their needs by expanding on the training content; and (b) stakeholder groups familiar with the internal operations of the training institute highlighted the need for improved organizational operations, including stronger relationships with external partners, clearer communication processes, and more clarity around organizational goals and priorities. Overall, staff and leaders of the training institute placed high value and trust in the report's findings as evidenced by (a) the second author's team being immediately hired to conduct the evaluation in Year 2, (b) the report being shared with key stakeholders and funders, (c) staff and leaders openly discussing what they learned through the evaluation (educative authenticity), and (d) most of the recommendations made in the report being implemented in Year 2 (catalytic authenticity).

Case Discussion

In our experience with REAM, we have learned several important lessons concerning how to conduct a credible and feasible rapid evaluation. First, it is critical to have clarity and

Table 4
Example Data Analysis Matrix

Evaluation Questions

Informant ID	Stakeholder Group	Perceived Impacts of Training[a]	Perceived Strengths of Training Institute	What Needs to Be Improved
O23	Training participant	<u>Strengthened external community relationships:</u> • Our organization now has strong community relationships. <u>Strengthened board of directors:</u> • Board membership continues to grow (three new members added).	<u>Experiential learning model:</u> • I love the fact that we got to practice the skills in the session. <u>Quality of instructors:</u> • Teachers were qualified experts in their field.	<u>Expand communication or marketing efforts:</u> • Too few people know about this institute. If they had a Web site, more would come.
034	Training participant	<u>Expanded orgs volunteer base:</u> • More volunteers and more are taking on leadership roles. <u>Increased fiscal viability:</u> • Have been more successful at grant applications. Have increased revenues by 30%.	<u>Experiential learning model:</u> • It was great to have the assignments that made us practice the skills at home before the next sessions.	<u>Expand training offerings:</u> • Need more trainings for orgs that are just forming.
051	Funder	<u>Strengthening external relationships:</u> • Participating orgs seem to be expanding the number of partners they have.	<u>Developing reputation as the best:</u> • Some are saying they offer the best training out there.	<u>Increase and diversify funding base:</u> • Need more funders from a variety of contexts.

a. Underlined text indicates a theme. Bulleted text is verbatim/evidence taken from interviews to illustrate theme.

agreement on the targeted evaluation questions during the initial weeks of the effort. There is little, if any, time to drastically shift the orientation of a rapid evaluation, thus sufficient time should be spent up front ensuring that the right questions are being asked of the right people.

Second, data collection and analyses must happen simultaneously, particularly when dealing with qualitative data. This is expedited when interview data are not transcribed. Digital audio recorders can serve as an excellent resource by allowing immediate access to the conversation via the media player on a computer, without having to spend the time and resources creating verbatim transcripts.

Third, report writing should begin as soon as possible in the process. Once "answers to the targeted questions" have emerged, evaluators can begin to outline their responses. This allows for more reflective writing time at the end of the process, increasing the quality of the final report.

Fourth, rapid data collection efforts are more successful when they do not compete with other rapid demands targeted informants are experiencing. For example, in Year 2 of this project, data collection efforts competed with a new reporting deadline that many of our informants were facing. Because these reports were mandated (and quite time consuming), and participation in our interviews was optional, we lost many potential informants (because of time conflicts) and needed to identify and recruit others to interview. Because this delayed the completion of data collection, we ultimately needed to renegotiate a new deadline for the final report. This suggests that rapid evaluators should work closely with their collaborative partners to identify any unforeseen conflicts that might impede the successful implementation of a project; sometimes, timelines may need to be renegotiated.

Fifth, rapid evaluation requires a team of skilled and experienced evaluators; these projects provide little time for training or for learning a new evaluation methodology. Thus, as evaluators positioned within academia, we only pursue rapid evaluation projects when we can create an experienced team.

Finally, we have found that that confirmatory and dependability audit trail processes recommended by Guba and Lincoln (1989) were critical for ensuring that methods are logical and strategically employed and that findings accurately portray stakeholder perceptions.

Conclusion

Our review of the literature on REAM reveals a family of methods employed in a vast range of contexts that have provided quick, reliable assessments of responses to humanitarian crises (RTE), HIV/AIDS primary care programs (RFE), patterns of HIV/STI transmission (RA), health care systems (REM), and management of local natural and human resources (PRA). Despite their range of topical and geographical applications, these approaches share a common core of methods and principles: mixed methods coupled to a rapid, iterative, team-based approach.

A central challenge that all REAM users face is achieving a balance between speed and trustworthiness. We have seen that these approaches are indeed quick, but what can be done during such a short, intense period of time to also ensure their credibility and trustworthiness? As we illustrated above, we have found Guba and Lincoln's (1989) Fourth Generation Evaluation well suited for guiding and judging the adequacy of a REAM project. Although Guba and Lincoln's model assumes specific ontological, epistemological, and methodological approaches (i.e., it follows a constructivist paradigm and promotes qualitative methods), its emphasis on criteria that judge whether an evaluation's process and findings accurately portray stakeholders' experiences and concerns (Guba & Lincoln, 1989) seems particularly relevant to many REAM studies. Of

course, as Table 3 illustrates, some of their criteria are difficult to follow in a rapid evaluation project. We particularly struggled with our inability to pursue prolonged engagement and obtain adequate member checks (from trainees) of our data. These are techniques that we, as qualitative researchers, often rely on to determine the trustworthiness of our findings. However, what is valuable about Guba and Lincoln's framework is that it provides a multitude of strategies for each criterion, allowing evaluators to adopt practices appropriate for and feasible within any given project.

Regardless of the framework used to judge a REAM effort, however, the strength of any project depends not on the particular methods used, but how they are used and the conscientiousness of the investigators. Have they described the context of discovery in sufficient detail that similar contexts of application can be readily identified (transferability)? Were the methods of research employed in a systematic and consistent fashion over time, and is this documented (dependability)? Can final conclusions be traced back first to the data and second to the sources (confirmability)? Like any research method, then, when the techniques of REAM are employed carefully, systematically, and conscientiously, they are more likely to produce findings worthy of trust.

When, then, does it make sense to use REAM? The compressed timeline of REAM suggests the appropriate context of application: REAM is rapid because there is typically an urgent need for information on the basis of which action is to be taken. Whether findings will be used to improve responses to humanitarian crises or design a more effective tuberculosis control program, the use of REAM results is almost always instrumental. Rapid evaluation and assessment method is not designed to contribute to an existing body of theory but to provide information of sufficient quality at key decision points to improve the quality of decision making and, by extension, the effectiveness of actions subsequently taken. Given this, a necessary condition for a successful REAM project is an evaluation consumer who is both open to learning and ready to take action on the basis of findings. Also, REAM is more easily implemented in settings where conflict or disagreement across stakeholders is low. In these contexts, the time it takes to develop appropriate understandings and adequately address the differences across stakeholders runs counter to the need for quick, timely information.

Note

1. Grigori Potemkin is a legendary figure who is reputed to have built impressive fake villages along a route that Catherine the Great was to travel.

References

Anker, M., Guidotti, R., Orzeszyna, S., Sapirie, S., & Thuriax, M. (1993). Rapid evaluation methods (REM) of health services performance: Methodological observations. *Bulletin of the World Health Organization, 71*(1), 15-21.

Aral, S., St. Lawrence. J., Dyatlov, R., & Kozlov, A. (2005). Commercial sex work, drug use, and sexually transmitted infections in St. Petersburg, Russia. *Social Science & Medicine, 60*, 2181-2190.

Argyris, C., Putnam, R., & Smith, D. (1985). *Action science: Concepts, methods, and skills for research and intervention*. San Francisco: Jossey-Bass.

Bartsch, D., & Belgacem, N. (2004). *Real-Time Evaluation of UNHCR's response to the emergency in Chad*. Geneva, Switzerland: Evaluation and Policy Analysis Unit, United Nations High Commissioner for Refugees.

Beebe, J. (2001). *Rapid assessment process: An introduction*. Walnut Creek, CA: Alta Mira Press.

Bellavita, C., Wholey, J., & Abramson, M. (1986). Performance-oriented evaluation: Prospects for the future. In J. Wholey, M. Abramson, & C. Bellavita (Eds.), *Performance and credibility: Developing excellence in public and nonprofit organizations*. Lexington, MA: Lexington Books.

Bentley, M., Pelto, G., Straus, W., Schumann, D., Adegbola, C., de la Pena, E., et al. (1988). Rapid ethnographic assessment: Applications in a diarrhea management program. *Social Science & Medicine, 27*(1), 107-116.

Chambers, R. (1994a). The origins and practice of participatory rural appraisal. *World Development, 22*(7), 953-969.

Chambers, R. (1994b). Participatory rural appraisal (PRA): Analysis of experience. *World Development, 22*(9), 1253-1268.

Desmond, N., Allen, C., Clift, S., Justine, B., Mzugu, J., Plummer, M., et al. (2005). A typology of groups at risk of HIV/STI in a gold mining town in north-western Tanzania. *Social Science & Medicine, 60,* 1739-1749.

Garrett, J., & Downen, J. (2002). Strengthening rapid assessments in urban areas: Lessons from Bangladesh and Tanzania. *Human Organization, 61*(4), 314-327.

Greenwood, D., Whyte, W., & Harkavy, I. (1993). Participatory action research as a process and as a goal. *Human Relations, 46*, 175-191.

Guba, E., & Lincoln, Y. (1989). *Fourth Generation Evaluation*. Newbury Park, CA: Sage.

Guerrero, M., Morrow, R., Calva, J., Ortega-Gallegos, H., Weller, S., Ruiz-Palacios, G., et al. (1999). Rapid ethno-graphic assessment of breastfeeding practices in periurban Mexico City. *Bulletin of the World Health Organization, 77*(4), 323-330.

Jamal, A., & Crisp, J. (2002). *Real-Time Humanitarian Evaluations: Some frequently asked questions*. Geneva, Switzerland: Evaluation and Policy Analysis Unit, United Nations High Commissioner for Refugees.

Kirsch, H. (1995). The use of rapid assessment procedures: Focus groups and small-scale surveys for community programs. In H. Kirsch (Ed.), *Drug lessons and education programs in developing countries*. New Brunswick, NJ: Transaction Publishers.

Kresno, S. Harrison, G., Sutrisna, B., & Reingold, A. (1994). Acute respiratory illness in children under 5 years in Indramayu, West Java, Indonesia: A rapid ethnographic assessment. *Medical Anthropology, 15*, 425-434.

Lee, T. & Price, M. (1995). Indicators and research methods for rapid assessment of a tuberculosis control pro-gramme: Case study of a rural area in South Africa. *Tubercle and Lung Disease, 76*(5), 441-9.

Lewin, K. (1948). *Resolving social conflicts: Selected papers on group dynamics* (G. W. Lewin, Ed.). New York: Harper & Row.

McNall, M., Welch, V., Ruh, K., Mildner, C., & Soto, T. (2004). The use of rapid-feedback evaluation methods to improve the retention rates of an HIV/AIDS healthcare intervention. *Evaluation and Program Planning, 27*, 287-294.

Miles, M., & Huberman, A. (1994). *Qualitative data analysis: An expanded sourcebook*. Thousand Oaks, CA: Sage.

Needle, R., Trotter, R., Singer, M., Bates, C., Page, G., Metzger, D., et al. (2003). Rapid assessment of the HIV/AIDS crisis in racial and ethnic minority communities: An approach for timely community interventions. *American Journal of Public Health, 93*(6), 970-979.

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

Rifkin, S. (1992). Rapid appraisals for health: An overview. *Rapid Rural Appraisal Notes, 16*, 7-12.

Rutta, E., Williams, H., Mwansasu, A., Mung'ong'o, F., Burke, H., Gongo, R., et al. (2005). *Disasters, 29*(4), 291-309.

Sandison, P. (2003). *Desk review of Real-Time Evaluation experience*. New York: United Nations Children's Fund.

Scrimshaw, N., & Gleason, G. (Eds.). (1992). *Rapid assessment procedures: Qualitative methodologies for planning and evaluation of health-related programs*. Boston: International Nutrition Foundation for Developing Countries.

Sonnichsen, R. (2000). *High Impact Internal Evaluation: A practitioner's guide to evaluation and consulting inside organizations*. Thousand Oaks, CA: Sage.

Trotter, R., & Singer, M. (2005). Rapid assessment strategies for public health: Promise and problems. In E. Trickett & W. Pequegnat (Eds.), *Community intervention and AIDS*. New York: Oxford University Press.

United Nations High Commissioner for Refugees. (2002). Real-time evaluations. *Forced Migration Review 14*, 43.

Utarini, A., Winkvist, A., & Ulfa, F. (2003). Rapid assessment procedures of malaria in low endemic countries: Community perceptions in Jepara district, Indonesia. *Social Science & Medicine 56*, 701-712.

Vincent, N., Allsop, S., & Shoobridge, J. (2000). The use of rapid assessment methodology (RAM) for investigating illicit drug use: A South Australian experience. *Drug and Alcohol Review 19*, 419-426.

Wholey, J. (1983). *Evaluation and effective public management*. Boston: Little, Brown.