



EVALUATIVE THINKING

1. EVALUATIVE THINKING AND OBSERVATIONS

How Can Observations be used Effectively?

Observations are conducted to view and hear actual program activities. The purpose of conducting observations is to describe the program thoroughly and carefully, and in sufficient detail so that users of the observation report will know what has occurred and how it has occurred. Observations involve **looking and listening**. A particular strength of observations is that data are collected in the field, where the action is, as it happens.

Observations can be focused on programs overall, components of programs or participants in programs. They are used to both add clarity to other data and as a strategy to directly obtain data about program implementation and/or participant outcomes as they happen. Observations can also be used to determine changes over time.

The process of observing, like all other types of data collection, can affect regular program operations. The evaluator must be able to monitor the effects and take them into consideration when analyzing data (e.g., session leaders may be extremely nervous when their activity is observed, participants may behave better or worse). **It is always important for the evaluator to recognize s/he is observing only a subset of possible actions and to determine by asking and reviewing program/activity descriptions, how typical a session is.** What happens in one session or to selected participants may be very different from what happens in another session or to other participants. Observations of special or unique activities, such as culminating sessions or performances can provide an opportunity to focus specifically on outcomes.

Conducting observations multiple times, by multiple trained observers, of multiple subjects (e.g., regular program sessions, final or culminating sessions, participants and staff, actions and interactions) and in conjunction with other types of data collection (i.e., surveys, interviews or record reviews) adds rigor to program evaluation. Like surveys and interviews, good observations require a good protocol, administration plan and analysis plan, all developed before the observations are conducted.

A note about inter-rater consistency. Observation, perhaps more so than other strategies requires particularly careful forethought regarding data collection. Description is critical, but the observer often also needs to categorize information during an observation. Observers must know what they are seeing when they see it. If more than one observer is collecting data about the same subject, it is especially important to ensure they will categorize it similarly (i.e., that there is inter-rater consistency). Training and preparation are critical.

Observation Benefits

1. Trained observers get direct experience with a program which allows them to see things first hand rather than piecing things together from other descriptions or responses.
2. Trained observers have the opportunity to see things that may routinely escape awareness among program participants.
3. Trained observers have the opportunity to learn about things that the program participants may be unwilling or unable to talk about in an interview.
4. Observations permit observers to move beyond the selective perceptions of others, and to incorporate their own perspectives.
5. The impressions and feelings of observers can be added to the observation data and the observers' knowledge and direct experiences can be used as resources to aid in assessing the subject program.

Reminder: While observers' reactions, or how they feel about a subject can contribute to understanding of findings, observer's reactions must be recorded separately from actual observation descriptions. It is the observer's primary job to record what they see and hear, not what they think is happening.

What Methodological Decisions Must be Made Before Conducting Observations?

Observations are less commonly used in program evaluation than surveys or interviews because they are time consuming to conduct and analyze. But there are multiple possible observation approaches and they can be especially useful for collecting outcome data that is difficult to capture through surveys or interviews. Effective data collectors always determine the purpose for the observation, collect background information about the subject(s), and decide the following before conducting observations.

- *What will be observed and how the observation protocol (see following) will be structured. Specifically, what will the observation focus on (settings, individuals, skills, program or program component delivery, interactions, decision-making, best or promising practices)?*
- *How to choose what to see (i.e., what is the sampling strategy?). [See note below.]*
- *Whether to observe a 'performance' by the subject or to attend one or more regular sessions, or both.*
- *Whether the observer's presence will be known, or unannounced? Who should know? How much information about the purpose of the observation will be disclosed and to whom? [See box page 6.]*
- *How much detail will be sought (can a checklist observation do the job)?*
- *How long and how often will the observations be?*

Sampling: Selecting Observation Subjects

Because qualitative data are usually collected for much smaller numbers of subjects, sampling strategies used to determine sample composition for surveys are not relevant for selecting what gets observed. While interview respondents are typically selected to comprise **Purposeful Samples** which include information-rich cases (i.e., those that are typical, extreme or deviant, that can provide maximum variation or are considered critical cases), observation "sampling" has fewer rules of thumb. Common strategies include the following:

- **If the goal is to document program implementation, use of space or regular use of specific practices**, single observations can be conducted. They are always best conducted at "typical" times for the program.
- **If the goal is to characterize program delivery and/or participant response to programs in general**, observe at least two regular sessions for the duration of the sessions. These should be spaced apart depending on the schedule of the program, with an initial observation conducted close to the beginning of service delivery, and a second or follow-up observation scheduled near the conclusion of the program cycle.
- **If the goal is to document participant outcomes**, select one or more sessions where participants are expected to demonstrate learning or skills (such as during a culminating activity, or during a skills test administration).
- **If the goal is to collect rich, comprehensive information about the way a program is delivered, how participants respond and change in response to the program, and whether specific practices are consistently used**, multiple observations must be conducted.

Caution: Program schedules shift. Ascertain that the sessions for observation will be happening as planned. Be sure to review the subject selection strategy with stakeholders so they will trust observation summary results.

What's off limits? It is unethical to observe confidential sessions (such as a counseling or case management visits) or participant interaction unless the observer has specific permission to do so. Observers should limit their participation in session activities.

Comprehensive Program/Session Observation Protocol EXAMPLE

Program Name: _____ Observer's Name: _____

Date: _____ Time Observation Began: _____ Time Ended: _____

Before the observation begins, briefly describe for #1 below, what you expect to be observing and why you have selected it.

1. Subject of the Observation.

At the very beginning of the observation, describe the setting. Be sure to note any changes as the observation proceeds.

2. Describe the **program setting** (color, size, shape, number of desks/tables, number of windows, furniture or equipment in the space room, temperature, noise level)
3. Describe **how the session begins** (who is present, what exactly is said to initiate)
4. Describe the **chronology of events** in regular intervals (e.g., every 15 minutes for 1 hour session)
5. Describe the **interactions that take place during the observation**. (*Record who is interacting and how, and any changes during the observation. For example note that younger students only interact with other younger students until the experiments are conducted then interactions happen between older and younger students. Be particularly aware of interactions that involve the main focus of the evaluation – i.e., who or what you are observing.*)
6. Describe how **decisions** are made during the observation period, who makes them and how are they communicated? Document examples of decisions that are made during the observation. (*Be sure to record who is making the decision.*)
7. Describe **Nonverbal communication** (How do participants get attention? How much do they fidget, move around? How do participants: dress, express affection, physically place themselves in the setting?)
8. **Describe program activities and participant behaviors.**
9. How did participants respond or react to what was happening in the program during the observation? **Roughly what proportion (some, most, all) are actively engaged?**
10. **How does the program end?** (What are the signals that the activity is ending? Who is present, what is said, how do participants react, how is the completion of this activity related to other activities?)

Note the above protocol example has had all space removed to facilitate placement in this article. A standard observation protocol can be a list to use, or can include blank space so observations can be recorded directly on the protocol.

What Happens During and *Immediately After* Observations?

- Use the observation protocol to guide the observation and record data about what is happening for the selected subject(s) and to complete check-lists (see page 5).
- *Review observation notes and make clarifications where necessary. Be sure to clarify any abbreviations used in notes and elaborate on details if necessary. [Transcribe if feasible or necessary.]*
- *Evaluate the results of the observation: determine whether there were any barriers to observation; if follow-up is needed.*
- *If multiple observers are used to collect data, compare results as soon as possible to determine if there is sufficient inter-rater consistency (i.e., to determine if multiple observers recorded information similarly about the subject).*

Conducting observations requires skill and additional decision-making. Like interviews, first and foremost, the observer must be familiar with and understand all sections of the protocol so record keeping will be complete. Review and practice (including mock sessions where possible) are strongly advised. Before the observation begins, the observer should plan where s/he will be stationed to conduct the observation and how much s/he will participate in any activities. The observer should be prepared to clarify the purpose of the observation and what will happen to the data.

Analyzing and Reporting Observation Data

After all planned observations have been conducted, the observer must summarize key findings. Summarization of observation data is used to describe strategies or participants and/or to highlight or help clarify certain findings determined through surveys, interviews or record review. Generally, summary statements are made (e.g., most of the sessions were productive) and "*snippets*" or excerpts from field observation notes are included. Note that observation data are different from survey and from interview data, (although check-list observation data are similar to close-ended survey responses, see following) therefore analytical strategies are different for them.

Where analyses of survey data tend to focus on frequencies and cross-tabulation (i.e., using numbers), and interview data tend to focus on themes and quotes, analyses of observation data tend to focus on descriptions (i.e., words and pictures). Just as it is common when analyzing interview data, to count words, or ideas or themes and present them in numeric terms, observations can be summarized numerically as well, especially if multiple observations have taken place, or if multiple subjects were observed. Collecting, analyzing and reporting about data from observations is an excellent way to get comprehensive findings and it is definitely advisable if one is evaluating an unfamiliar subject. Observations like interviews are labor intensive, however, and only result in rich data about a fairly narrowly framed set of findings. Combining data collected through observations with data collected via surveys, interviews or record reviews can enhance coverage.

Analyzed Observation Data Example from an Arts Education Program

- *Many different types of arts activities were undertaken, and personal development segments were integrated with them. Of the 57 different combinations of programming at the 10 sites, all but three covered the planned curricula. During each observation, all but a few students were following directions.*
- *At all sites, ongoing projects were underway and examples of participant work were readily visible. Teaching artists were demonstrating skills, giving youth opportunities to try the skills, and providing one-on-one assistance as needed.*



Excerpt Example: Training Setting and Activity (*Field notes have been adjusted for reporting*)

There was a maximum of 15 participants in the room, and a minimum of 9 participants, about 13 stayed for most of the training (some left or were there intermittently throughout the session). At least 5 different people uninvolved in the training passed through the room. The room was set between the front entrance and a back computer lab and staff offices. At least 6 participants arrived late (after the first topic had been covered) and two departed early (before most of the training had been completed).

During the discussion portions of the training, participants are talking about verbal triggers they have experienced (e.g., being whistled at or propositioned after exiting the subway). During the exercise, one participant walked across the room and the other participants verbally offered triggering harassing comments. A majority of the group contributed comments and then added their own recollections of similar circumstances. The volunteer described what it felt like to be "harassed" during the session, and multiple participants added examples of that happening to them and how they responded.

Using Observation Check-lists

Check-lists, or sections of observation protocols that include checklists are a hybrid between a survey and an observation. The observer is responding to a survey about the subject using categories and ratings that can easily be summarized, (including whether something is present or not). Check-list type observations are especially effective when the observer is determining whether a particular skill or desired action is being demonstrated. When “best” or promising practices are known, they can easily be integrated into an observation protocol and substantiated with additional descriptions from the observation. For example, using what is known about productive literacy development strategies, a protocol like the following could be used to collect data about rooms used for literacy development activities.

LITERACY DEVELOPMENT: PROGRAM FACILITIES/EQUIPMENT								
In the space below, provide a brief description of the site including size of the space (sq. footage, # of rooms), arrangement of furniture, distinguishing features.								
Is the program site accessible to all potential participants?..... <input type="checkbox"/> No <input type="checkbox"/> Yes								
Does the site have an area where participants can casually interact (hang out)? <input type="checkbox"/> No <input type="checkbox"/> Yes								
Is participant work displayed at the site? <input type="checkbox"/> No <input type="checkbox"/> Yes								
Are there any unmet maintenance needs? <input type="checkbox"/> No <input type="checkbox"/> Yes								
Please rate the following features of the physical environment at the site.								
1 = Poor	2 = Fair	3 = Good	4 = Excellent	1	2	3	4	NA
Attractiveness of physical facility (freshly painted, good lighting etc.)				<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Condition, appropriateness, and quantity of furniture				<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ability to accommodate both large and small group activities at same time				<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Availability of independent study areas				<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Availability of reading areas with comfortable seating				<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Availability of computers for participant use				<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Participant access to the internet				<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Availability of books or other reading materials				<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Availability of art supplies for participant projects				<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Attractiveness to children				<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Overall facility rating				<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Check-lists can also be prepared to record observations of skills and abilities. These are best developed in conjunction with program providers. They can be both developed and conducted by program staff and then shared with evaluators, or conducted by both program staff and external evaluators who have been trained to recognize the skills in question. Skills and abilities can be observed at varying points in a program, or tests can be devised to determine if specific skills have been developed. Observations can be used to record both objective measures (e.g., whether a participant can perform a skill or answer a question correctly) and subjective ratings related to those measures or other expected actions (e.g., determining or judging the value of a performance or display). The following is an example of a summary of observation findings showing skill development.

Example of Summarized Check-List Observation Data on Outcomes

Table 1: Participant Vocabulary Development: Spring 2013, Observation Results

Participant Number	Followed Directions Using New Program- Specific Vocabulary	Number of Correctly Identified Items (n = 20)	Overall Performance Rating
Participant 1	Some Directions	12	Okay
Participant 2	All Directions	19	Excellent
Participant 3	Most Directions	15	Very good
Participant 4	All Directions	20	Excellent
Participant 5	Some Directions	14	Good
Participant 6	All Directions	19	Excellent
Participant 7	Most Directions	16	Very Good
Participant 8	Most Directions	20	Excellent
Participant 9	Some Directions	15	Good
TOTAL/Average	All or Most Directions = 6	16	Very Good/Excellent = 6

Note Table 1 displays summarized observation data. The protocol to collect these data would require checkboxes for each activity where directions were required as well as checkboxes for recording correct identification of each of the 20 terms, and a final assessment for the performance with directions for assigning the ratings *Ok*, *Good*, *Excellent*, *Not Good*. An observation protocol would have been completed for each participant 1 – 9.

A quick note about the role of observers during observed sessions:

Generally speaking, observers should not fully participate in the sessions they observe. However, to minimize disruptions sometimes it is useful if an observer sits or circulates among participants and responds to introductory activities or attempts some of the activities being observed. Participation ranges from no participation (being a “fly on the wall”) to fully partaking of the activities to get a direct experience. (The latter is often referred to as “going native” and is usually not done as it is very hard to be a participant and an observer at the same time.) Regardless of choices made about participation, observers should never share unanalyzed results of observations with either participants or leaders. Similarly session leaders should not attempt to conduct observations and lead the sessions as each action requires substantial effort by leaders.

What can you do to Enhance Evaluative Use of Observations?

There are definite steps that can be taken to enhance the effectiveness of observations so that findings from them can inform data-driven decisions. (*Using findings and making data-driven decisions are key elements of evaluative thinking.*)

1. PREPARE FOR OBSERVATIONS, TRAIN OBSERVERS

Good observation protocols are critical to conducting good observations. Make sure that all methodological decisions have been addressed (see page 2), that everyone who is expected to conduct observations understands how to use the protocol and is at least generally familiar with the subject(s). **Remind all observers that you can't observe everything that is happening all the time.** Check for inter-rater consistency regarding protocol use.

2. CAREFULLY SELECT WHICH AND HOW MANY SUBJECTS TO OBSERVE - BE SURE STAKEHOLDERS AGREE WITH CHOICES

3. DEVELOP AN ANALYSIS PLAN BEFORE CONDUCTING OBSERVATIONS, ADDRESS INTER-RATER CONSISTENCY CONCERNS

Planning for use of observation data before it is collected is an essential step. As with surveys and interviews, this will help strengthen the content of the observation protocol and help the observer(s) make choices about how much and what to focus on. **It should be possible to identify what will be done with the results of each observation question and checklist and if not, perhaps the questions should be deemed un-necessary.** Analysis plans also provide specific directions for what to do with data once it is collected so the analyst is not overwhelmed and the analysis process, which for observations largely entails cleaning up descriptions, looking for similarities and differences across descriptions (where available) and adding results of checklist data -- can be expedited.

A good analysis plan must specify what procedures will be conducted with each question or group of items on the protocol. It should also include directions for:

- how data will be disaggregated or partitioned if necessary (i.e., organized into meaningful segments -- e.g., by observation site or subject, or time of observation – initial, final etc).
- how data will be assessed (e.g., whether the description illustrates a positive or desired action or condition vs. an undesired action or condition).
- how to ensure that all involved in analysis agree on assessment of data
- how findings will be presented (narrative summaries, bulleted statements, excerpts from field notes, pictures or recordings of sample findings).

4. CONDUCT ANALYSES ACCORDING TO THE PLAN

Read/review completed sets of observations and record general and checklist summaries. Use field note excerpts to illustrate findings. Be sure to verify/validate your findings with key stakeholders (i.e., see if they interpret the findings similarly). Revise summaries and displays accordingly.