

The Performance Technologist's Toolbox: Observations

by Anne F. Marrelli, PhD, CPT

This second article in the Performance Technologist's Toolbox series focuses on the data collection method of observation. In observations, data are collected by systematically observing employees as they perform their jobs and documenting their behaviors. Well-conducted observations can provide much valuable qualitative and quantitative information about the employee, the job, and the work environment. They often yield data that cannot be obtained through paper- or conversation-based methods, such as questionnaires, interviews, or focus groups (Swanson, 1994).

There are six dimensions to consider in planning observations: overtness, standardization, degree of observer participation, time sampling, documentation, and level of inference.

Dimensions of Observations

Overtness

In overt observations, the employees who will be observed are informed in advance of the observations; they may also be told which behaviors will be observed and why. In some cases, the observer may give the employees directions about what to do during the observations, such as performing a specific task. Overt observations are often combined with interviews, in which the observer asks the employees to explain the reasons for their actions.

In covert observations, employees are not aware of some or all aspects of the observation, including the specific behaviors being observed or when, how, or by

whom they are being observed. In some covert observations, the observer is in view, but employees are not told what behaviors are being observed. The observer does not interact with the worker in any way but allows the employees to perform the normal tasks in the usual sequence (Swanson, 1994).

The "mystery shopper" is another type of covert observation often used to assess the job performance of salespeople or call service center representatives. In this approach, the observer poses as a regular customer or caller and observes how the employee handles the transaction. Employees typically know what behaviors are being observed, but they do not know when or by whom they will be observed. In fully covert observations, employees may be viewed at work through a two-way mirror or video camera and are not told that they are being observed.

Overt observations are most appropriate for situations in which employees are not likely to significantly alter their behaviors if they know they are being observed. Overt observations are also an effective approach for collecting data about knowledge workers, because often the observer will need to ask these workers questions to understand what they are doing and why. Covert observations are most appropriate when employees are likely to change their behaviors if they are aware of the observation.

In most cases, however, to maintain trust with employees, it is important to give them advance warning that their performance will be observed at an unspecified time and to explain how the information collected will be used.

Standardization

Observations may be made of a representative sample of routine situations or of a standardized situation set up for all employees being observed. The actual job situation yields more natural behavior, while the standardized approach controls for situational variables that could introduce contamination (Popham, 1975).

Participation

The participation of the observer in the activities of the employees being observed may range from none at all to full participation as a member of the group or team. The observer who is actively involved can develop a deeper understanding of the target behaviors and the variables that affect them, but objectivity usually suffers (Borg & Gall, 1979; Cozby, 1981).

Time

The observation periods need to be selected to obtain a representative sample of the target behaviors. For a very routine job in which the same simple tasks are repeated throughout the day, a one-hour observation of each subject may suffice. For complex jobs in which tasks vary by day, week, month, or season, observations may need to be scheduled over a period of several months.

There are several different approaches to determining when and how much of the behavior that occurs during an observation period will be recorded.

- **Continuous observation:** All behaviors of the employee that are demonstrated during the observation period are noted. This method usually does not focus on a specific set of behaviors and is often used in exploratory studies to identify key behaviors that will later be studied in more depth. The observer writes a protocol, a brief narrative of everything that happens and when it happens.
- **Frequency-count observation:** The observer notes each instance when specific target behaviors occur, for example, talking with a coworker. Other behaviors are not recorded. This technique is often applied in comparing the quantity or content of specific behaviors before and after a performance improvement intervention is implemented.
- **Interval recording:** The behavior of the subject is observed and recorded at pre-specified intervals, for example, every 5 minutes or the first 15 minutes of every half hour. Other behaviors are not recorded (Borg & Gall, 1979). The interval recording approach is especially useful when complex behavior is being observed and the observer needs time to record or evaluate the behavior between each segment of observation.

Documentation

Observations may be recorded in narrative form or by completing a standard form. The narrative record may consist of an

anecdotal record or critical incident in which only key events are documented, or a protocol in which all events are recorded. Standard forms often consist of behavior checklists in which the observer records the presence, absence, frequency, or duration of specified behavior. Sometimes the observer will be asked to record codes for different behaviors. Narrative records produce a fuller picture of work behavior than standard forms, but standardized forms are easier to complete, can be objectively scored, and produce more quantified data, for example, a count of the number of times a specific behavior occurred (Phillips, 1997; Popham, 1975; Worthen & Sanders, 1973).

Video cameras may be used to record behaviors during an observation. Videotapes offer the advantage of being repeatable; the researcher can play the videotapes as many times as necessary to fully review the target behavior. They are useful in observing complex behaviors that need to be carefully studied before inferences can be made and also when specially qualified reviewers need to observe the behavior but are not available for in-person observations (Borg & Gall, 1979). However, video recording often makes employees nervous and alters their normal behavior (Phillips, 1997).

Sometimes audio recordings are used to monitor the performance of employees who spend much time on the telephone with customers, such as salespeople or call center representatives. The procedure is typically explained to employees so they are aware that a certain number of their calls will be randomly monitored. Computers can also be used to “observe” employees who enter data or perform other keyboarding tasks. The computer can record the time it takes employees to complete a task, the sequence of steps they use, the number of errors they make, and so on (Phillips, 1997).

Inference

There are different levels of inference required of observers in recording the events observed.

- **Descriptive:** Observers record clearly observable behaviors such as the number of times a worker performs an activity, the use or non-use of job aids or tools, or the length of time required to complete a phone call.
- **Inferential:** Observers are asked to draw inferences about what a behavior represents, for example, the observer is asked to record instances in which a supervisor demonstrates supportive behavior to the employees supervised.
- **Evaluative:** Observers are asked to draw inferences and make judgments about behaviors, e.g., observers are asked to record instances in which a supervisor demonstrates supportive behavior to the employees supervised and rate the quality of the support provided (Borg & Gall, 1979; Popham, 1975).

The higher the level of inference, the less reliable the measurements typically are. The performance technologist is

often faced with a choice between measuring less-significant behaviors more effectively or more significant behaviors less effectively (Popham, 1975).

Applications of Observations in Performance Technology

Observations are a valuable tool for the performance technologist. Several applications are explained below.

Competency Modeling and Job Analysis

Observations can provide valuable preliminary information about a target job role because they offer a sample of what the job “looks like” in real life. They are especially valuable for routine jobs in which the incumbent performs the same tasks many times throughout the day because in a few hours the observer can see the full spectrum of job activity. Comparison samples of exemplary and poor performers are often observed to identify the competencies required for superior performance (Marrelli, 2001).

Performance Analysis

Observations can be a potent approach to identifying performance improvement needs. An astute observer will see the types of errors employees are making, the redundancies in processes, the work methods applied, the frustrations employees face, and the problems they experience. Observations will also provide information about interactions among employees and their supervisors, safety or housekeeping problems, and areas in which employees seem to lack needed information. Differences between expected and actual results can also be observed (Swanson, 1994).

Cause Analysis

Because observations yield a rich, full picture of the work environment, they are an ideal data collection method for identifying the underlying causes of performance problems. They help the performance technologist uncover obstacles to effective job performance as well as identify the workplace supports that facilitate superior performance. Observations are also useful in analyzing work flows and processes for inefficiencies and misalignment as well as identifying ineffective communication patterns. They also uncover differences in work behaviors among exemplary, average, and poor performers (Swanson, 1994).

Evaluation

Observations are an effective method of assessing employees’ mastery of the skills taught in training courses that cannot be reliably measured with written tests (Marrelli, 1998). Observations are also very effective in determining whether employees are applying on the job what they have learned in training. A wide range of other performance improvement interventions, including programs, processes, and

tools, can also be evaluated with observations. For example, in a situation where the physical layout of an office is redesigned to encourage more interaction among employees, observers can count the number of interactions before and after the redesign. In another example, a program designed to encourage more employees to participate actively in staff meetings, observations will provide a clear picture of the type and amount of participation before and after program implementation.

Individual Performance Assessment

Individual employee performance can be effectively assessed through careful observations. For example, the “mystery shopper” observation mentioned above is a valid approach to determining whether employees are meeting the performance standards established for interactions with customers. In another example, training managers often sit in on classroom sessions to observe the facilitation skill of instructors; the instructors subsequently receive feedback and coaching in how to improve instruction.

Advantages and Disadvantages of Observations

Observations as a method of collecting valid, reliable data offer various advantages and disadvantages.

Advantages

- Because actual job behavior is directly observed, the validity of observations is especially strong.
- Both verbal and nonverbal behavior can be observed.
- Interactions of job incumbents with colleagues, managers, and customers can be observed to create a full and rich perspective of behavior and the work environment not possible in other data collection methods.
- Many people in an organization trust observation data more than data collected via other methods because they see the observations as direct measures of actual work behavior. Observations can therefore facilitate buy-in for organizational initiatives.
- Observations can be adapted to many different situations.
- Observations can measure constructs such as racial prejudice through their demonstration in behavior that cannot be accurately measured with other data collection methods.

Disadvantages

- Observations are expensive and time consuming. Especially with more complex jobs, an observer may need to spend many hours over many days to develop a comprehensive understanding of the job. To ensure sound data and to compute reliability estimates, at least two independent observers need to evaluate the same situation.
- The quality of the data depends heavily on the skill, objectivity, and sensitivity of the observers. They need to be carefully trained to produce accurate reports. If the

observers are not properly trained, the data collected may be unreliable because of the biases of the observer, poor recording techniques, or rating errors.

- The presence of an observer may modify employees' typical behavior.
- Employees may object to being observed. If there are strong objections, they may purposely alter their usual behavior.
- It can be difficult to interpret the observations, especially when the work observed involves intangible tasks such as problem solving.

Guidelines for Planning and Conducting Observations

Planning Observations

The diligence invested in planning a systematic observation process is strongly related to the validity and reliability of the data collected. Several planning steps are needed:

- Determine what behavior will be observed.
- Develop forms for recording data.
- Select the observers.
- Prepare a schedule of the observations.
- Train observers in what and how to observe.
- Inform participants of the observations in advance (Phillips, 1997).

Obtaining data related to complex behavior that is objectively observable yet relevant requires careful thought. Limit the focus to the number of behaviors that observers can realistically attend to effectively and record; not all behaviors can be observed in a single observation. Once the target behaviors have been selected, it is essential to specifically define them so that all observers have a common understanding (Borg & Gall, 1979).

It is important to understand the cycle of work before scheduling observations to ensure the observers observe long enough and at the right times to obtain a representative sample of the target work or situation. Some tasks will be performed only at the beginning or end of the month, some times of day will be busier than others, etc. (Swanson, 1994).

Recording Data

Accurate recording of observations is essential because memory cannot be trusted (Pelto & Pelto, 1979; Swanson, 1994). Develop a standard observation form for all observers to use in recording their observations; use of the form will improve the consistency and objectivity of observations. Even if observers are to record events in narrative format, a standard structure for the critical incident or protocol will facilitate consistency and comprehensiveness.

Develop a draft version of the form, then pilot test it and revise it as needed. If a checklist is used, describe the behaviors in enough detail so that observers need only to check each behavior as it occurs. Attached instructions should specifically

define each behavior and provide examples and non-examples of each behavior. In evaluative observations, observers are required to evaluate behavior as it occurs. Do not attempt to obtain precise discriminations. Three-point scales are usually the limit of discrimination that can be made with a reasonable level of reliability. It is almost never advisable to use ratings finer than a five-point scale (Borg & Gall, 1979).

Training Observers

Observers will require extensive training. The following steps are suggested:

- Select observers carefully. Some people are sharp observers while others have difficulty noticing or remembering details about situations they observe.
- Meet with the observers to review the form to be used. Devote sufficient time to each item to allow the observers to develop a thorough understanding of the behavior to be observed and how it should be documented. Review examples and non-examples for each behavior.
- Show videotape recordings of situations similar to those to be observed. Each time a target behavior appears on the video, stop the tape and explain why the behavior falls within the definition of the behavior.
- Conduct practice sessions using the videotapes in which the observers record the behaviors using the standard forms. Review their documentation and provide detailed feedback. Resolve any inter-observer disagreements.
- Repeat the video practice sessions until the observers have reached a high level of skill in observing and recording and the desired level of agreement among raters has been reached.
- Conduct an additional practice session in a real-life situation as similar as possible to the actual observation situations. Ensure that the observers know how to work as unobtrusively as possible.
- Assess the skill of each observer before allowing them to conduct observations.
- Schedule the actual observations to begin within a day or two of the last practice session. If observations will extend over more than a week, conduct a weekly refresher training session or the data will become less reliable as the raters lose their common frame of reference (Borg & Gall, 1979; Marrelli, 2001; Phillips, 1997; Pelto & Pelto, 1978; Popham, 1975).

Conducting Observations

At least a few days before the observations, explain the purpose of the observations and the procedures the observers will use to the employees who will be observed. On the day of the first observations, allow time for the employees to adapt to the presence of the observer before actually beginning the observations (Cozby, 1981).

If the observers cannot determine what a work behavior is or what its purpose is, if the technology in use is unfamiliar, or if a controversy arises about how the work should be done,

the observers should ask the study director to resolve these issues before continuing to observe (Swanson, 1994).

Carefully review the observers' documentation as they complete it to identify any problem areas. Talk to the observers to obtain their reactions. Do the observers seem to understand how to use the recording forms appropriately? Are the reports thorough? Do the reports include the level of detail needed? Are the observers accurately capturing the target behaviors? Are the observers having any problems conducting the observations? Do the employees being observed seem to be behaving naturally? Are the time periods allotted for each observation sufficient to capture and record the target behaviors?

Complementary Data Collection Methods

Observations can be used as a preliminary data collection technique to develop an understanding of a target job or work situation and important behaviors. The results of the observations can then be confirmed in surveys or focus groups in which more people are involved. Observations can also be applied to probe more deeply into data collected through surveys, focus groups, or work logs.

Case Study

In the large county government agency where I began my career, my team and I developed written selection tests for protective services positions within the Sheriff, Fire, and Marshal's Departments. One test we developed was for the entry-level Sheriff's position, Deputy Sheriff Trainee. The first step in developing the examination was to conduct a job analysis. Observations were a key data collection method for the job analysis. To help us understand the job, we rode in patrol cars with Deputy Sheriffs, observing how they spent their work days. We watched them stopping drivers for traffic violations, frisking people they suspected of carrying illegal substances, cruising the streets looking for trouble, answering citizens' questions, giving directions, checking in at the patrol station, investigating auto accidents, talking with their buddies, and completing reports. While we drove around with the deputies, we asked them questions about their work. We also observed the current Deputy Sheriff Trainee class in the Sheriff's Training Academy. We watched the trainees marching in formation, doing calisthenics, practicing at the firing range, and sitting in a classroom learning about the laws they would later need to enforce. During and immediately after both the ride-along and academy observations, we recorded what we had seen in narrative form.

Other data collection methods we used in the job analysis included interviews, focus groups, review of reading materials distributed in the Academy, and a review of a large sample of the patrol reports the deputies had recently written. We used the data we obtained in the observations and the other data collection methods to identify the competencies required

for the Deputy Sheriff Trainee position. The content of the selection test we developed was based on these competencies.

Conclusion

Observations offer a versatile approach to collecting data in many different situations. Many of the nuances of work behavior that are missed in paper and interview-based techniques can be readily perceived in observations. The impact of actually seeing people at work often provides the performance technologist with a deeper level of understanding of the job, its context, and challenges than is possible with other methods. 🌱

References

- Borg, W.R., & Gall, M.D. (1979). *Educational research: An introduction* (3rd ed.). New York: Longman.
- Cozby, P.C. (1981). *Methods in behavioral research* (2nd ed.). Palo Alto, CA: Mayfield.
- Marrelli, A.F. (2001). *Introduction to competency modeling*. New York: American Express.
- Marrelli, A.F. (1998). Ten evaluation instruments for technical training. In Kirkpatrick, D.L. (Ed.), *Another look at evaluating training programs* (pp. 58-65). Alexandria, VA: American Society for Training and Development.
- Pelto, P.J., & Pelto, G.H. (1979). *Anthropological research: The structure of inquiry*. Cambridge, UK: Cambridge University Press.
- Phillips, J.J. (1997). *Return on investment in training and performance improvement programs*. Houston, TX: Gulf.
- Popham, W.J. (1975). *Educational evaluation*. Englewood Cliffs, NJ: Prentice-Hall.
- Swanson, R.A. (1994). *Analysis for improving performance: Tools for diagnosing organizations and documenting workplace expertise*. San Francisco: Berrett-Koehler.
- Worthen, B.R., & Sanders, J.R. (1973). *Educational evaluation: Theory and practice*. Worthington, OH: Charles Jones.

Anne Marrelli, PhD, CPT, has more than 20 years of experience in the performance technology field with deep and broad expertise in selection and assessment, performance management, instructional design and development, performance improvement, evaluation, organizational development, competency modeling, and leadership development. Her former employers include Caliber Associates, American Express, Hughes Electronics, Educational Testing Service, and the County of Los Angeles. Anne earned a BS degree in Psychology and Sociology from Whittier College and MS and PhD degrees in Educational Psychology from the University of Southern California. She is the author of numerous journal articles, book chapters, and technical reports. Anne may be reached at annemarrelli@earthlink.net.