



SOME EVALUATION WORDS

Kate Lorig, DrPH
kate@selfmanagementresource.com
www.selfmanagementresource.com

Many special terms are used to describe evaluations. These allow communication but sometimes also are used as a secret language or as a means of intimidation. The important thing to know is that there is no one right way to conduct evaluation. Rather, good evaluations are part art and part science. Even the very best evaluators are learning new things every day. The following are some common evaluation terms. Being familiar and comfortable with them should serve you well.

Process or Formative Evaluations

Process or *formative evaluations* ask questions about how your program is operating. In other words, you are evaluating the process? These evaluations can be very simple or quite complex. Process questions include: How many people attended? Where did people find out about the program? Were some classes better attended than others? What was the dropout rate? Were the people who dropped out like the people who stayed, or were they different? How did the people who participated differ from the general population? Process evaluations can also be used to find out if the program is being implemented according to a set protocol and if the instructors are liked by the participants.

Outcome or Summative Evaluations

Outcome or *summative evaluations* ask whether your program is doing what you wanted it to do. Most commonly, outcome evaluations ask questions about changes in behaviors, health status, or health care utilization. However, sometimes you may put on a program with the expressed purpose of improving patient satisfaction. If this is the *purpose* of your program, then satisfaction could be an outcome, although it is usually a measure of process. The key consideration is the *intended outcome* of your program and whether you are evaluating this. Of course, if you are not clear about what you want a program to do, it is very hard to evaluate outcome.

Quantitative Evaluation

Quantitative evaluation collects data that are easily converted into numbers. It can be used for either a process or an outcome evaluation. Examples of quantitative evaluations are:

- Asking participants to rate instructors on set scales
- Using instruments to measure depression, satisfaction, disability, etc.
- Using clinical data such as blood pressure or blood glucose
- Using chart or self-report data on utilization, such as the number of days in hospital or the number of emergency room visits

Quantitative evaluations are best used when you are very clear about the question that you want answered – that is, when you have a hypothesis. If you have a clear hypothesis, then almost anything can be quantified.

Qualitative Evaluation

Qualitative evaluation collects data that are not easily converted into numbers. (Even qualitative data can often be converted to numbers, but this takes thought and skill.) Qualitative data are almost always words. They are collected by means of semi-structured interviews, focus groups, listening to the conversation of others, open-ended questions, or observation.

Qualitative evaluations are best when you are not at all clear about what is happening or when you do not know exactly what questions to ask. These evaluations are often used to form a question or a hypothesis for a quantitative study.

To make things even more complicated, both quantitative and qualitative methods can be used in the same evaluation. For example, you might ask a series of questions about depression, health status, and behaviors, and then ask an open-ended question about what participants found most useful from an intervention.

Objective/Subjective

Objective and *subjective* are words that are often used to confuse or intimidate the beginning evaluator. In theory, *objective* refers to anything that can be verified by a standard test and outside observation. Such data as blood pressure, cholesterol levels, and number of visits to a doctor are considered objective. *Subjective* refers to any data that can be biased by or are the opinion of the reporter. Subjective data are considered not to be valid or "real" and are therefore suspect.

The problem is that the line between objective and subjective is not at all clear. Let us examine self-report of visits to physicians (considered by some to be subjective) as contrasted with chart audit (considered by some to be objective). It is true that an individual may over- or under-report the number of visits or may confuse a visit to a lab with a visit to a doctor. However, a chart audit may also have problems. The individual may see several doctors at different places, and not all the charts may be audited, leading to underreporting of visits. The person doing the chart audit may count as only one visit seeing a primary care physician and then being referred to a specialist on the same day. Thus, chart audits are not always valid.

Other examples of objective data are blood pressure and cholesterol level. Many people have been declared hypertensive on the basis of blood pressure readings taken in a physician's office. In fact, the blood pressure may be high because of the anxiety surrounding a doctor visit. As for cholesterol level, laboratory data sent to five different labs have a good chance of getting several different cholesterol readings, some as much as 20 to 30 points apart.

The point is that seemingly objective data can be subject to bias. At the same time, subjective data can be valid. For example, pain and stress are always considered subjective. The important issue is not objective versus subjective, but rather whether the data are valid. Do not ever let anyone intimidate you by questioning the "objectivity" of your data. Be prepared to explain why you believe that the data are *valid*.

Valid

Valid is the key word when collecting any type of data. For data to be valid, they must meet two tests. First, you must get the same answers if you do the same test or ask the same questions twice within a short period of time. This is called *reliability*. However, something can be reliable without being true. For example, a clock is 15 minutes fast. Five people look at the clock and report the same time. The data are reliable, but they are not correct. For something to be valid, it must be reliable and correct.

To determine the correctness or validity of data, we usually compare it with some "gold standard". For example, we have an individual self-report his or her disability and then have a physical therapist (PT) give the person a series of tests and they report the disability. If the two reports are the same, then the self-reported data are usually considered valid because they are the same as the "gold standard" or the rating of the PT.