

A 5-step method for identifying performance measures for any program in 45 minutes

The following five step scripted process is the best way to select the most important performance measures and identify a Data Development Agenda for any program or service. With practice, this process can be completed in about 45 minutes. Participants should each have a copy of the performance measurement summary on page 28.

Step 1. How much did we do? Draw the four quadrants on a piece of flip chart paper. Start in the upper left quadrant. Write down the measure "number of customers served." Ask if there are better, more specific ways to count customers or important subcategories of customers and list them, such as the number of children with disabilities served.

Next, ask what activities are performed. Convert each activity into a measure. The activity of "training people" becomes number of people trained. Paving roads becomes number of miles of road paved. When you're finished, ask if there are any major activities that are not listed. Don't try to get every last detail, just the most important categories of customers and activities.

Step 2. How well did we do it? Ask people to review the common measures listed in the upper right quadrant of the performance measurement summary. Write each one that applies in the upper right quadrant of the flipchart paper.

Next take each activity listed in the upper left quadrant and ask what measures tell how well that particular activity was performed. If you get blank looks, ask if timeliness or accuracy matters. Convert each answer into a measure and be specific. The timeliness of case reviews becomes percent of case reviews completed on time. If you are not sure whether a measure goes in the upper right or lower right quadrant, put it where you think best and move on. All the measures in both quadrants will be considered equally in Steps 4 and 5.

Step 3. Is anyone better off? Ask "If your program works really well, in what ways are your customers' lives better? How could we observe this? How could we measure it?" Create pairs of measures (number and percentage) for each answer. For example, the number of clients who get jobs goes in the lower left quadrant. And the percent of clients who get jobs goes in the lower right quadrant. It saves time, when entering these measures, to write them only once in the lower right quadrant, and place # signs in the lower left quadrant across from each measure.

Identifying whether anyone is better off is the most interesting and challenging part of this process. Dig deep into the different ways in which service benefits show up in the lives of the people served. Explore each of the four categories of better-offness: Skills / Knowledge, Attitude / Opinion, Behavior, and Circumstance. If people get stuck, try the reverse question: "If your service was terrible, how would it show up in the lives of your customers?"

Look first for data that is already collected. Then be creative about things that could be counted and how the data could be generated. It is not always necessary to have data for all of your customers. Data based on samples can be used. Pre and post testing can be used to show improvement over time in skills, knowledge, attitude and opinion. When no other data is available, ask clients to self-report about improvements or benefits.

Keep in mind that all data have two incarnations: a lay definition and a technical definition. The lay definition is something that everyone can understand. The technical definition gives the exact way in which the measure is constructed. For example, “high school graduation rate” is a lay definition with many possible technical definitions. The easiest technical definition is the number who graduate on June 15th as a percentage of enrollment on June 1st. This will always be close to 100%. A tougher technical definition would compare graduation numbers to enrollment on September 30 of the previous year. A still tougher definition would compare graduation to the enrollment of 9th graders four years earlier. Each technical definition constitutes a separate measure.

When you complete step 3, you will have filled in the four quadrants with as many entries as possible. In steps 4 and 5, we use a shortcut method to assess the communication, proxy and data power of each measure and winnow these down to the most important measures.

Step 4. Headline measures: Review the list of upper right and lower right quadrant measures and identify those for which there is good data. By good data we mean that timely and reliable data for the measure is available today or could be produced with little effort. Put a circle next to each one of these measures. Next, ask "If you had to talk about the performance of your program in a public setting, such as a public hearing or conference, and you could only use one of the measures with a circle, which one would you choose?" Put a “#1” by the answer. Then ask "If you could have a second measure... and a third?" You should identify no more than 3 to 5 measures. These should be a mix of upper right and lower right measures. These choices represent a working list of headline measures for the program.

Step 5. Data Development Agenda: Ask, "If you could buy one of the measures for which you don't have data, which one would it be?" The word “buy” is used because data is expensive both in terms of money and worker time. With a different colored marker, write DDA #1 next to the chosen measure. "If you could buy a second measure... and a third?" List no more than 3 to 5 measures. These measures are the program’s Data Development Agenda *in priority order*.

This process leads to a three part list of performance measures:

Headline measures: Those 3 to 5 most important measures for which you have good data, the measures you would use to present your program's performance in a public setting.

Secondary measures: All other measures for which you now have good data. These measures will be used to help manage the program, and will often figure in the story behind the baselines.

Data Development Agenda: A prioritized list of measures where you need new or better data. You will later need to make a judgment about how far down this list you can afford to go.