Building an Evaluation Activity

Tobacco Control Evaluation Center

When developing the work plan and choosing activities, it is helpful to go about it strategically. Start with the end goal of the objective in mind and then work backwards, building in and connecting the activities necessary to achieve the objective.

To facilitate the plan-writing process, CTCP provides sample activities known as wizards (writing templates) with typical wording for various activities. The activity wizards must be adapted to specific project goals, needs, and capacity. Activity parameters need to be tailored so they are appropriate for the intended population. Though the details and order may differ, each activity should cover:

- **Intended outcomes:** purpose, topic/focus, audience/end-user
- **Parameters:** sampling, population of interest, recruitment/training, mode, instrument, waves, logistics (timing, location, personnel)
- **Analysis and action plan:** links to intervention activities or other strategies, explains how results will be used
- **Reciprocity:** a dissemination plan to share results back with the population from whom data are collected

The following structure can be used for writing a new evaluation activity description:

> To [**PURPOSE**: inform, improve, measure, confirm, monitor] [**END-USER**: policy makers, staff, coalition, store/MUH owners, general public, etc.] about [**TOPIC/FOCUS:** what you want to know, document, or accomplish] a [**METHOD:** survey, observation, interview, focus group, media record, policy record/document review, Photovoice, litter audit, describe other evaluation activity] will be conducted with **[SAMPLE SIZE** #-#] of [**POPULATION OF INTEREST** (unit of analysis, identifying characteristics, inclusion/exclusion criteria)]. A [**SAMPLING METHOD**: purposive, random, stratified, cluster, convenience, census] sample will be selected/recruited from [constituency (board of supervisors, city council, community leaders, store owners, chamber of commerce, church leaders, influencers, thought leaders, school leadership, general public, location type, etc.)] at/in/from [**LOCATION** as applicable]. A [**MODE**: paper and pen, online, mobile, verbal, visual, interview guide] **INSTRUMENT** and protocol will be developed or adapted [in consultation with TCEC or from another source] pilot tested, revised, and tailored for use with the intended population and cultures. This [pre, post, or pre and post] measurement will be conducted in in [#] **WAVES** before and after [intervention, e.g., policy implementation, educational campaign, community outreach, etc.] for [duration: # of minutes, rounds, groups]. [**PERSONNEL**: data collector pool e.g., staff, coalition, youth, volunteers] will **be recruited, trained, and assessed for readiness** [e.g., practice, knowledge tests, monitoring] to administer instrument protocols. Data will be **ANALYZED** using [descriptive statistics, inferential statistics, content analysis] and results will be disseminated with [**STAKEHOLDERS**: specific staff, staff in general, coalition, policy makers, general public, etc.] in appropriate [**FORMATS**: presentation, fact sheet, summary report, media release, podcast, etc.]. Lessons learned will be used to [**ACTION PLAN**: measure change over time, support or inform next steps, link to interventions as appropriate, etc.].

Example: To improve staff understanding of community knowledge, attitudes, and perceptions regarding smoke-free parks, a public intercept survey will be conducted with a convenience sample of 150-300 visitors of parks in Savannah using a mobile device. The survey and protocol from the previous work plan with be used. This pre/post measurement will be conducted in two waves before and one wave after smoke-free parks policy implementation. Surveys will be conducted by project staff who will be trained and assessed for readiness with during trainings to administer instrument protocols. Data will be analyzed using descriptive statistics and inferential statistics to document support/opposition to potential policy strategies, knowledge, awareness, beliefs, and demographic information provided by survey participants. Results will be shared with program staff and other stakeholders such as the coalition and potential partners to inform next steps and improve interventions.

More considerations for building evaluation activities

After being submitted, every proposed work plan is reviewed and scored by a team of program, budget, and evaluation specialists. They are looking to see that each plan consists of a rational series of activities that work toward achieving the objective. The logic of the number, sample size, and use of evaluation activities is a principal indication of the overall quality of the plan. So, it is important to get this right!

An effective work plan relies upon not only the appropriate mix of evaluation activities, but also the proper number, scope, timing, and size of activities. Every evaluation decision should be driven by utility. Reviewers are looking for the answer to these questions: What purpose will it serve? How will the data collected inform or support interconnected activities and progress toward the objective overall?

It is important to avoid the mistake of thinking that cramming in more evaluation activities makes the plan more rigorous or fund worthy. Each activity needs to add value and have a purpose. Otherwise, it is just busy work.

So how can a plan writer tell if an activity is necessary? Employ end-use strategizing thinking: What information or evidence would be helpful to know and have to achieve the objective? Typically, it is crucial to conduct formative research first through policy record review, media activity record review, key informant interviews with policy makers, and observation data. Then, collect and measure evidence of the issue and of support to show decision makers, increase public knowledge and awareness of the issue and potential solutions, measuring readiness to take action, and timing of intervention and evaluation activities.

Also consider whether it would be advantageous to be able to collect similar data points in waves over time or from multiple sources/types (e.g., qualitative AND quantitative) in order to triangulate or confirm the validity of findings.

Sample Sizes

When determining an effective sample size for an evaluation activity, think about what threshold will be compelling, convincing, and representative enough to inform and/or persuade critical audiences. For example, will survey results collected from just 100 city residents persuade a city council member that her constituents support a particular policy direction? In a small community, it might be, but in a large community, probably not.

A sample size calculator will show the ideal size paired with confidence intervals. [See https://www.surveymonkey.com/mp/sample-size-calculator/]. With a smaller or more contained population, aim to get 100% (called a census of the population, rather than a sample or portion of it). For example:

- Coalition Satisfaction Survey must include a census of all members
- Education/Participant Survey for data collectors must include a census of all trainees
- Key Informant Interviews should include a minimum of 5 or a census
- Young Adult Tobacco Purchase Survey should include a minimum of 25 stores or a census of all tobacco retailers
- Focus Groups should include a minimum of 2 groups with 6-10 adults or 5-7 youth

Also remember to specify the breakdown by jurisdiction in the activity description. For example, state, "The sample size will be 25-35 in the 2 jurisdictions and 2 comparison communities for a total of 100-140."

To make data compelling for a policy maker, it may be important to survey residents of that policy maker's district. Instead of a sample size for a city of 100,000, it might be a sample size for a district of 20,000. Getting more specific about sample sizes and audiences will get better quality data and easier-to-manage data collection efforts.

Quantitative data collection activities such as surveys or observations generally aim for some level of representativeness meaning that who or whatever is part of the sample includes the breadth and depth of the makeup and characteristics of the whole population of people or things being measured.

Qualitative data collection activities do not typically aim for representativeness and extrapolation of results to the larger population. The samples are often selected purposively rather than randomly and can therefore be fairly small.

Documenting the rationale for sampling choices and incorporating these details into the activity description can help staff team recall the sampling decisions when it comes time to actually carry out the evaluation activity. For example, "In order to ensure that the primarily Hispanic/Latino voices are included in the evaluation, 4-5 focus group discussions will be conducted with approximately 6-10 people in each group. Two of the groups will be comprised of (female) Latina tenants, two groups of (male) Latino tenants and one group containing a mix of Latinx tenants."

There is a variety of different sampling strategies to choose from: simple random, stratified random, cluster, purposive, convenience. Be sure that the strategy is a good fit for the evaluation activity and the sample composition. In addition, the plan should define the inclusion and exclusion criteria which specifies the characteristics of what should be included or not included in the sample.

For more details on sample sizes, sampling method, and other sampling decisions, see https://tobaccoeval.ucdavis.edu/parameters-and-sampling

Budget Allocations for Evaluation

The number and type of evaluation activities and their sample sizes will be dependent upon how much staff, resources, and time the project will have available to conduct the activities. Think about the value and cost that different types of

activities will require. Key informant interviews can be conducted by one or two people, while store or park observations may require teams of data collectors and data collector trainings.

Some activities benefit from the more costly but well-honed skills of an evaluator or facilitator; others can be carried out by trained volunteers and a team leader who monitors quality control throughout the data collection process.

Percent Deliverables for Evaluation

When assigning the percent deliverable for an activity, estimate the amount of staff or contractor time it will take to plan, conduct, analyze, and write up results, given the data collection method, location, and sample size.

Do a reality check by converting that deliverable into a dollar figure (x percent of the funding amount). Does the amount seem reasonable for the work output? For example, a cost of $40,000 to conduct 4-5 key informant interviews will be sure to raise a red flag for plan reviewers. Experienced evaluators are likely to have more realistic projections and will know how long certain tasks will take to complete.