International Center for Alcohol Policies
Analysis. Balance. Partnership.

ICAP TOOLKIT

A Guide to Evaluating Prevention Programs

(DECEMBER 2010)

channel
research

This Toolkit is intended for those working in the beverage alcohol industry—producers, trade associations, and social aspects organizations (SAOs)—who are currently undertaking or planning to undertake programs to prevent alcohol-related harm.

The Toolkit provides an overview of what is involved in evaluating prevention programs, laying out the necessary steps and identifying different available options.

Annex A is a stepwise guide to how this Toolkit may be applied to evaluating two specific examples of initiatives: an awareness campaign around the use of designated drivers and a school-based alcohol education program.

Annex B offers links to existing guides on evaluation and to examples of successful programs, including alcohol education, screening and brief intervention, alcohol-impaired driving, and responsible hospitality.

Table of Contents

1. Introduction: Why Are Evaluations Important?

Targeted prevention measures around alcohol provide information, raise knowledge and awareness, and aim to help change behavior in an effort to reduce potential harms around drinking. They focus on the three elements of drinking patterns:[1]

- "at-risk" individuals
- risky behaviors
- high-risk contexts and settings

These measures include alcohol education, drink-drive countermeasures, prevention of underage drinking, training for those who sell and serve alcohol, initiatives to reduce violence and crime—particularly around licensed premises and the nighttime economy—and interventions aimed at identifying and modifying harmful drinking.

Companies that produce beverage alcohol, their trade associations, and social aspects organizations (SAOs) have invested heavily in such prevention efforts. In many cases, these interventions have been implemented with considerable success. However, this success remains largely undocumented. As a result, measures like responsibility messages, alcohol education, or server training are often convenient targets for advocates of greater regulation around beverage alcohol.

What is lacking is an *evidence base* around the effectiveness of many targeted interventions. This evidence base can only be achieved through proper evaluation.

This Toolkit offers guidance on conducting an evaluation—from planning to collecting data and interpreting the findings. Its Annexes provide step-by-step examples of evaluating two specific types of interventions likely to be implemented by industry stakeholders (an awareness campaign around the use of designated drivers and a school-based education program) and offer links to further reading.

---

1   Stimson, G., Grant, M., Choquet, M., & Garrison, P. (Eds.). (2007). Drinking in context : Patterns, interventions, and partnerships. New York: Routledge. Available: www.icap.org/Publications/DrinkingInContext

International Center for Alcohol Policies (ICAP). (2008). Guide to creating integrative alcohol policies. ICAP Policy Guides Series, available at: www.icap.org/PolicyTools/ICAPPolicyGuides

International Center for Alcohol Policies (ICAP). (2008). Quick reference guide to the ICAP Blue Book: Implementing alcohol policy and targeted interventions. ICAP Policy Guides Series, available at: www.icap.org/PolicyTools/ICAPPolicyGuides

1.1 Dispelling Some Evaluation Myths

There is a belief among some that evaluating programs may not be the most efficient use of time and resources. Five common myths surrounding evaluations are worth debunking:

- ***Myth 1: It is sufficient to implement a program without evaluating it.***

  WRONG. Programs that have not been evaluated do not carry much weight. Evaluation is the only way to demonstrate that a particular approach or initiative has been successful. Nothing dispels the notion that education, server training, and other efforts are simply window dressing better than demonstrating plainly and clearly with facts and figures that industry-supported initiatives can be as effective as any others, and that they are held to the same standards. This is not to say that unevaluated programs do not help society, but lack of evaluation invites claims that the effort is not worthwhile and deprives program sponsors and researchers of data that could help development of future interventions.

- ***Myth 2: It is better to spend resources on running initiatives than on evaluation.***

  WRONG. Although resources are scarce, and it may be tempting to focus exclusively on implementation, no intervention is complete without proper evaluation. A single well-evaluated intervention that can be supported by evidence as to its effectiveness can be more valuable than several initiatives that have no evaluation. The latter may leave potential donors guessing as to whether the initiatives worked or what lessons can be learned for developing new programs.

- ***Myth 3: Since there is already evidence to show that different types of initiatives (e.g., drink-drive countermeasures, social norms campaigns, and brief interventions) can be effective, there is no need to evaluate each new program individually.***

  WRONG. Evidence of past success is a strong argument in favor of choosing a particular approach or type of intervention. However, just because a certain approach has worked in the past does not mean that it will work again: Social context, conditions, target groups, and many other factors are likely to be different in each case. Therefore, any intervention should be evaluated in its own right—this is the only way to demonstrate whether it has actually worked in the particular setting.

- ***Myth 4: Evaluation is too complicated and requires the involvement of outside experts.***

  WRONG. Evaluating a program simply means asking the right questions and gathering the information in a structured way. This can be done by an outside expert, but there are also simpler evaluations that can be conducted internally. The choice of which evaluation to use depends on several considerations, including available resources and the complexity of the issue at hand. This is discussed in **Section 2.2: Types of Evaluation**.

- ***Myth 5: Evaluation only shows whether an intervention has succeeded or failed.***

  WRONG. While determining success or failure is certainly one of the main purposes of evaluation, many other things can be learned. An evaluation allows prevention efforts to be *improved* so that they can be implemented more effectively in the future. If an evaluation is framed correctly and asks the right questions, it can uncover a lot of additional and valuable information (see **Section 2.1: The Purpose of Evaluation**).

2. What Is Evaluation?

Evaluation offers a way to determine whether an initiative has been worthwhile in terms of delivering what was intended and expected. However, good evaluation can also answer other important questions.

> **KEY TERMS:**
>
> - *Input—describes the resources (human and financial) expended on the initiative.*
>
>   Included here are the human (number of people involved and time spent) and financial resources that have gone into developing and implementing a particular initiative.
>
> - *Output—refers to the tangible products developed for the initiative.*
>
>   For example, informational literature or pamphlets that may have been developed, TV spots that may have been aired, or services (such as counseling) that may have been provided for use during the initiative.
>
> - *Outcome—describes the achievements of an initiative and its immediate or direct effects on those who participated in it.*
>
>   For example, do the participants of a particular campaign know more than before? Have their attitudes changed? Can any immediate effects on their behavior be observed?
>
> - *Impact—looks beyond the immediate results of an initiative and identifies longer-term effects, as well as any unintended or unanticipated consequences.*
>
>   For example, an impact evaluation might examine whether an intervention's immediate positive effects on behavior were sustained over time.

2.1 The Purpose of Evaluation

Evaluations of prevention programs fulfill a number of functions:

1. Measure the program's outcomes and impact
   - Did the program achieve its stated objectives?
   - Did it reach its intended audience?
   - Was the size of the outcome as expected?
   - Did the program have unexpected or unintended consequences?
   - Are outcomes consistent with those of similar programs?

2. Inform future program planning and design
   - What are the strengths and weaknesses of a given approach?
   - What implementation problems have emerged?
   - Are measurement criteria appropriate and adequate?
   - Are confounding influences affecting outcomes (e.g., other interventions that may have been aimed at the same issue or target group)?
   - Have new ideas emerged, and can they be tested?

3. Provide important internal lessons for those conducting programs

   For example, evaluations can offer feedback on whether the expenditure of financial and human resources needed for the program was justifiable:
   - Were funds used properly?
   - Is there a return on investment?

4. Ensure transparency and accountability

   Particularly where outside funding has been used on an initiative, evaluations help provide justification for the project. They can also be used as a form of stakeholder engagement, helping to gain buy-in from local community members, local authorities, and target audiences.
   - Are suitable systems in place to ensure sound financial reporting, monitoring, etc.?
   - Have lessons been taken on board for future initiatives?

5. Provide broader lessons about good practice
   - What lessons can be learned from this approach?
   - Are there lessons about policy options?
   - Do the results support existing evidence?

> **QUESTIONS for Evaluation:**
>
> 1) Has the intervention worked? How well?
> 2) Can the design and performance be improved?
> 3) Is expenditure on implementation justifiable?
> 4) Have lessons been taken on board for future programs?
> 5) Can this intervention contribute to good practice?

2.2 Types of Evaluation

Most evaluations fall into one of three categories:

- **process-based**
- **outcomes-based**
- **impact-based**

The choice of the most appropriate type of evaluation is guided by several factors, including the availability of resources and whether the evaluation is needed for internal or external purposes (see **Section 3.1: Who Should Evaluate?**):

- Process-based evaluations are useful in assessing how an intervention is being implemented or whether it is producing the necessary measurements.
- Outcomes-based and impact-based evaluations are best for tracking the results of an intervention.
- Process assessment is likely to be useful internally, whereas the focus on outcomes and impact can help justify the intervention both internally and externally.

Whichever evaluation model is used, data need to be collected in a systematic manner. Data may be

- **quantitative** (e.g., counting the number of drink-driving fatalities or the percentage awareness of a risk)
- **qualitative** (e.g., recording subjective views on whether a program has changed perceptions)

Successful evaluations often blend quantitative and qualitative data collection since there is usually more than one way to answer any given question (see **Section 3.3: Data Collection**).

> **KEY TERMS:**
>
> - **Qualitative data**
>
>   Descriptive and subjective.
>
> - **Quantitative data**
>
>   Measurable and definable in absolute, numerical terms. Can be subjected to statistical analysis.

2.2.1 Process-based Evaluations

Process-based evaluations are used to understand how a program works and delivers its results. They assess the activities that are being implemented and the materials that are used.

Process-based evaluations are intended to answer some of the following questions:

- What is required to deliver the program in terms of resources, products, and services?
- How are individuals implementing the intervention trained?
- How are participants selected and recruited?
- What are considered the program's strengths/weaknesses?
- What is the feedback from participants/partners about the implementation of the program?

2.2.2 Outcomes-based Evaluations

Outcomes-based evaluations are used to measure any changes immediately after program implementation and to establish that these changes have occurred in *response* to the intervention being evaluated.

Outcomes-based evaluations focus on the following questions:

- Which outcomes are being measured and why? (e.g., behavior change or change in knowledge or awareness)
- How will these outcomes be measured, specifically?
- What is the desired proportion of participants who will have undergone a change as a result of the intervention? Has this number been reached?

To be successful, outcomes-based interventions require the following:

- detailed information on the indicators that can be used to measure the desired outcomes (the best indicators are those that can be verified from administrative databases, surveys, third-party reports, or official statistics—e.g., the number of individuals participating in the program)
- a thorough assessment of how best to gather the necessary information—in other words, which **methodology** to use (see **Section 3.3: Data Collection**)
- a reliable and rigorous method for **analyzing and reporting findings** (see **4. Data Analysis and Interpretation**).

2.2.3 Impact-based Evaluations

By far the most complex and difficult to carry out, the impact-based evaluations examine the long-term effects of an intervention on participants:

- The most successful type of impact-based evaluation tracks effects **over extended periods of time**, rather than simply examining conditions immediately "before" and "after" the intervention has been implemented.

- Impact-based interventions can be further enhanced by including a **"control" or comparison group** against which to measure the "exposed" group (i.e., the one that has received an intervention) (see **Section 3.3: Data Collection**).

- Unfortunately, there can be **"confounding" contributors** to long-term "before" and "after" changes, aside from the program being evaluated.

Impact-based evaluation also requires information about the conditions before the intervention was implemented.

- For example, conducting an initiative aimed at alcohol-impaired driving will require that statistics and information be available about general drink-drive crashes and fatalities to provide a context.

3. Undertaking Evaluations

Most well-designed programs include provisions for evaluation from the very outset. Data collection, feedback, and measurement of impact and outcomes are built in.

TABLE 1 demonstrates how an evaluation can be integrated into the program's structure, listing the decisions and steps that must be made before, during, and after an intervention.

3.1 Who Should Evaluate?

Before addressing how each of the evaluation phases can be implemented, some thought needs to be given to who should conduct the evaluation.

There are four options regarding how an evaluation should be undetaken:

- **Internal** (i.e., carried out by those who are also implementing the intervention)

TABLE 1. Key Steps for Evaluating an intervention

Before the Intervention

Planning an Evaluation
- Identify the purpose of evaluation.
- Determine the appropriate approach to evaluation, depending on its purpose and available resources.
- Clearly define resources available and resources needed. This will determine whether evaluation will be conducted internally or by external experts, as well as its duration and follow-up.
- Develop evaluation methodology by identifying:
  - Target group(s) for evaluation results
  - Parameters for data collection (e.g., sample size, timeline for data collection and follow-up)
  - Appropriate questions

Data Collection
- Gather baseline data for future comparison (in outcomes- and impact-based evaluations).

During the Intervention

Data Collection (continued)
- Track how widely materials have been distributed and assess expenditure of time and resources.
- Gather information about intervention recipients.

After the Intervention

Data Collection (continued)
- Conduct post-intervention data collection to compare with baseline.
- Where appropriate, conduct a second wave of data collection at a later point to assess whether intervention effects persist over time and whether changes can be sustained.

Data Analysis and Interpretation
- Analyze data and compare with baseline.
- Interpret evaluation findings.
- Identify implications of findings.
- Report and disseminate evaluation findings.

- **Internally managed, with input from external researchers or organizations** brought in specifically for this purpose
- **Externally led and internally supported**
- **Independent external evaluation**

Each of the four approaches has different implications with regard to costs and time required to carry them out, as summarized in **TABLE 2**.

All four options have merit, and each has its own strengths and drawbacks. However, for some programs that deserve the expense investment, a case can be made that the most compelling and credible evaluations are those that are externally conducted.

3.2 Planning an Evaluation

Thorough planning helps design the steps taken during the evaluation process so that they are most likely to produce informative results.

It is important to begin planning an evaluation at the same time that the program itself is being developed (see TABLE 1).

The following should be identified and clearly defined during this stage:

- The **purpose of evaluation** (see **Section 2.2: Types of Evaluation**)

TABLE 2. Evaluation Type: Resource Implications

Credibility	Type of evaluation	Cost	Time[2]	Considerations
+	**Internal evaluation**	$	++++	• Good practice in evaluation. • Internal data collection and reporting. • Generates useful lessons and insights.
	Internally managed, with external input	$$	+++	• Better practice in evaluation. • When done right, will more likely be positively viewed in an external audit. • Results can be shared with wide audience or group of stakeholders, especially when rigor has been validated.
	Externally led, internally supported	$$$	++	• Generally considered best practice in evaluations. • Best for reporting to widest audience or stakeholders.
++++	**Independent external evaluation**	$$$$	+	• Adds credibility and objective assessment. • Will more likely be positively viewed in an external audit.

---

2   Time investment directly required of those implementing the program.

- Its **stakeholders**, whether external or internal (e.g., those who may be involved in the evaluation or be recipients of its results)
  - **External stakeholders** can include the intervention's target audience, partners involved in developing and delivering the program, community organizations and local actors interested in a given topic, and the media
  - **Internal stakeholders** can include funders, Board of Directors, or governance bodies
- Evaluation **methodology**, setting the parameters for data collection and appropriate questions (further discussed in **Section 3.3: Data Collection**)
- **Evaluation plan** and **terms of reference**, outlining key actions at various stages of both the evaluation process and program implementation

The following 8 principles can be integrated into the planning of most programs and are helpful in setting the stage for successful evaluations:

1. Do your homework.

What similar interventions have been done? Have they been successful and how have they been carried out? This provides validation for the program and the evaluation approach that may be taken.

2. Decide what needs to be addressed by the intervention.

A program may be designed to raise knowledge and awareness, build particular skills, or change certain behaviors. Knowing what the intervention is intended to influence or change will also identify what is to be measured and evaluated.

3. Define the program's target audience.

Is this program aimed at a particular group, for example, young people? If so, what age? It may be useful to focus on students as a convenient way of administering the program, if appropriate. In this case, evaluation should include only those who were enrolled both at the start and at the end of the initiative.

The target number of subjects that the intervention is meant to reach should also be identified. Ideally, the sample should be large enough so that the results are still useful even if some participants drop out.

4. Identify the program's key objectives.

Objectives should be clear, easily measurable, and realistic (i.e., there should be a degree of confidence that they can be achieved); they should also be identified at the very outset—during planning and before implementation of the program.

5. Identify the best approach to achieving the program's objectives.

This will depend to a large extent on the target audience, local conditions, and resources available. For example, if the literacy rate among the target audience is low, an intervention that relies mostly on written materials is probably not well suited for the given setting.

6. Set up clear criteria for measuring whether the program's objectives have been met.

How the objectives are measured will depend on what is most appropriate for the program. Where objectives are unrealistic or cannot be measured, the impact of the program cannot be demonstrated even if the program is worthwhile.

7. Determine how best to track progress and uptake of the program.

It is important to determine which method of **collecting** and **analyzing** data will be most useful and feasible. In some cases, particularly where large numbers of respondents need to be reached, surveys may be most appropriate. In others, interviews or focus groups are more useful (see **Section 3.3: Data Collection**).

8. Build in the ability to modify aspects of the program that may not be working.

Clear evaluation criteria and goals will help in the event that some mid-course corrections to the program are needed. Identifying what works and what doesn't is also useful if an intervention is to be repeated, increasing the likelihood of future success.

Once these questions are addressed, the evaluation goal(s), target audience(s), and available resources will become clear. All this will determine the appropriate approach to evaluation.

Creating an Evaluation Plan is a useful next step. Table 3 may be a helpful guide during this process.

An evaluation plan can help to

- Identify tasks and deliverables for each evaluation phase
- Assign roles and responsibilities
- Assess available and required resources
- Create a timetable for measuring progress

TABLE 3. Example of an Evaluation Plan Template

Phase	Tasks	Deliverables	Dates	Resources	Responsible
Planning & Inception					
Data Collection					
Reporting & Dissemination					

Preparing **Terms of Reference** for the evaluation at the planning phase can be useful. These can serve as a formal agreement of what will be done, regardless of whether the evaluation is conducted internally or externally. Terms of Reference should include:

- The purpose and timing of the evaluation
- The key questions to be asked
- The requirements of the evaluation team (if conducted externally)
- Expected structure of the report that will present evaluation findings
- Budget

Depending on the selected evaluation approach, baseline data for future comparison may need to be gathered at this early stage. **Section 3.3: Data Collection** provides further information on different approaches to data collection.

EVALUATION CHECKLIST: Planning

Key questions to consider while planning an evaluation:

- What is the purpose of your evaluation, and what do you hope to gain from it?
- Who are they key stakeholders you need to engage throughout the process?
- Who will conduct the evaluation?
- If it is an external consultant, who will be responsible within your organization for facilitating the process?
- Have you developed an evaluation plan, which outlines key tasks and deliverables and assigns roles and responsibilities?
- Have you decided on the questions to be asked during evaluation, chosen indicators, and identified data-collection methods?
- Have you decided on the structure of the evaluation report, and who will work on which sections?
- How will the evaluation findings be disseminated?

3.3 Data Collection

The success of the data collection phase often depends on the quality of the planning stage. How long data collection will take depends on the type of evidence required and the methods chosen for evaluation. The following questions will guide work at this stage:

- Where is the information to be found?
- Who is best positioned to get it and when?
- In what form should the information be stored, analyzed, and presented?

Two different types of data can be collected for any assessment:

- *Qualitative data* are descriptive and cannot be measured in absolute terms. They can be obtained from respondents' verbal answers to interview questions, focus group discussions, or written commentaries and responses to open-ended questions.
- *Quantitative data* are measurable and definable and can be converted into numbers and statistics. They are useful in showing absolute differences in what is being measured, such as percent changed. These data are derived from ratings, rankings, or "yes" and "no" answers to questionnaires.

Both types of data are useful. Where possible, it is often helpful to supplement objective, quantitative measures with more subjective and descriptive qualitative data.

How **data are collected influences whether they will be qualitative or quantitative.**

- Several approaches can be taken to gather data, ranging from simple to complex, and all can be used in evaluation.
- Each approach has its own set of strengths and weaknesses.
- The choice of the most appropriate evaluation approach will depend on an assessment of what is available and feasible on a case-by-case basis.

When data are collected, it is useful to also have additional information that is not directly related to the intervention being evaluated. For example:

- It is important to know **the sample size**, **the number of people who were reached**, and their **characteristics** (e.g., age, gender, socioeconomic status, etc.)
- In measuring change attributed to a particular intervention, it may be necessary to **establish the baseline** before the program is carried out (e.g., it is helpful to have the "before program" statistics, where available, on a particular topic).

The following are some of the most commonly used approaches to collecting data: **surveys**, **observations**, **interviews**, and **focus groups**.

EVALUATION CHECKLIST: Reliability and Consistency

Data collected must be reliable and consistent across different time points. Therefore, it is important to pay attention to the following:

- All sets of data should be collected in the same way, using the same set of instruments (e.g., questionnaire) at each time point.
- Where one-on-one interviews or focus groups are used, the same respondents should be included, to the extent possible, in each data collection round.
- Where this is impossible or when broader surveys are used, an effort should be made to maintain consistency among respondents (e.g., using the same target groups).
- Where feasible and appropriate, data can be collected from a control group who have not been targeted by the intervention but are similar to the intervention group in key characteristics.

3.3.1 Surveys

Surveys rely on data collection through questionnaires. These may be written (distributed in hard copy or electronically) or administered orally (e.g., by telephone).

Usually, in order to evaluate the impact of an intervention or a program, the survey is administered *twice:* once before and once after the intervention.

- Administering the survey *before* the intervention establishes **baseline responses**.
- A second round of surveys among the same group *after* the intervention will show whether there is a **change in knowledge, behavior, or whatever other outcome being measured**.

- In some cases, the survey may be applied a **third time to assess longer-term impact**. This is particularly useful when measuring behavior changes. It also helps to determine whether any short-term changes are sustained over time.

Another approach is to survey those who receive an intervention and those who do not, and to compare the results.

- This provides a "control" group for measuring the effect of a particular intervention.

Although surveys are used extensively in evaluation and, when constructed properly, can be very useful, the response rate to a survey may pose a challenge.

- Not all those who receive a survey will wish to respond to it. This may hinder the reliability and validity of the information that is extracted. It is therefore important to include a large sample to control for dropouts.
- Knowing the characteristics of the sample (e.g., age, gender, education level) will also help determine if the dropouts have some other significance.
  - For example, conducting a household telephone survey during weekday mornings or afternoons will miss all those who are at work. This is a useful piece of information to have when interpreting results.

3.3.2 Observations

Observation of individuals who have been exposed to an intervention (e.g., observing serving practices at retail establishments after a server training program or observing group dynamics during the delivery of a program) can help measure any changes in behavior or outcomes.

- As in surveys, a "before" and "after" assessment is needed to compare and measure effects.

Observation can help with determining whether a program is being delivered and implemented as planned and enable the evaluator to understand the situation and context.

However, observation is time-consuming and expensive, and care must be taken with interpreting its results, usually requiring an expert.

3.3.3 Interviews

This approach allows a one-on-one relationship between the evaluator and the respondents, and is particularly useful where personal contact is important.

- Interviews are especially helpful when the topic is complex and requires additional explanation or when

there are language (e.g., high illiteracy rate) or cultural barriers.

- Interviews also allow for immediate follow-up on interesting issues that may come up during discussion.

Given the nature of interviews, the sample size is likely to be relatively small. and the information obtained will depend heavily on the skill of the interviewer.

One useful form of interview involves so-called **key informants** who can speak on behalf of a larger population sample.

- **Key informants are experts in a particular field who can provide a broad view, representing the state of knowledge or different cultural views.**
- These can be academics with expertise in a particular field or members of a community with knowledge of and ability to speak for the community as a whole.

Information can be obtained from key informants informally or in more formal ways through interviews, hearings, or surveys. Key informants are frequently used to inform the work of bodies like the World Health Organization (WHO).

3.3.4 Focus groups

Focus groups combine elements of both observation and interviewing:

- A focus group is an interview with a gathering of 8 to 12 people, but uses group interaction to generate data and insights that would be unlikely to emerge in individual interviews.
- Originally used as a market research tool to learn the appeal of various products, the focus group method has been adopted in other fields as a way to gather data on a given topic.

Focus groups participants must be selected with care so that they are representative of the wider sample of individuals who were intended to be reached by the program.

Focus groups are often used in the pilot phase of a study to provide an idea of likely outcomes, followed up with a larger-scale approach to evaluation using, for example, a survey.

4. Data Analysis and Interpretation

The purpose of the data analysis and interpretation phase is to transform the data collected into credible evidence about the development of the intervention and its performance.

Analysis can help answer some key questions:

- Has the program made a difference?
- How big is this difference or change in knowledge, attitudes, or behavior?

This process usually includes the following steps[3]:

- Organizing the data for analysis (data preparation)
- Describing the data
- Interpreting the data (assessing the findings against the adopted evaluation criteria)

Where **quantitative data** have been collected, statistical analysis can:

- help measure the **degree of change** that has taken place
- allow an assessment to be made about **the consistency of data**

Where **qualitative data** have been collected, interpretation is more difficult.

- Here, it is important to group similar responses into categories and identify common patterns that can help derive meaning from what may seem unrelated and diffuse responses.
- This is particularly important when trying to assess the outcomes of focus groups and interviews.

It may be helpful to use several of the following 5 evaluation criteria as the basis for organizing and analyzing data:

- **Relevance:** Does the intervention address an existing need? (Were the outcomes achieved aligned to current priorities in prevention? Is the outcome the best one for the target group—e.g., did the program take place in the area or the kind of setting where exposure is the greatest?)
- **Effectiveness:** Did the intervention achieve what it was set out to achieve?
- **Efficiency:** Did the intervention achieve maximum results with given resources?
- **Results/Impact:** Have there been any changes in the target group as a result of the intervention?
- **Sustainability:** Will the outcomes continue after the intervention has ceased?

Particularly in outcomes-based and impact-based evaluations, the focus on impact and sustainability can be further refined by aligning data around the intervention's

- **Extent:** How many of the key stakeholders identified were eventually covered, and to what degree have they absorbed the outcome of the program? Were the optimal groups/people involved in the program?
- **Duration:** Was the project's timing appropriate? Did it last long enough? Was the repetition of the project's components (if done) useful? Were the outcomes sustainable?

In interpreting the results of evaluation, the issues of **Association, Causation, and Confounding** should be kept in mind.

4.1 Association, Causation, and Confounding

One of the most important issues in interpreting research findings is understanding how outcomes relate to the intervention that is being evaluated. This involves making the distinction between *association* and *causation* and the role that can be played by *confounding factors* in skewing the evidence.

4.1.1 Association

An association exists when one event is more likely to occur because another event has taken place.

However, although the two events may be associated, one does not necessarily cause the other; the second event can still occur independently of the first.

- For example, some research supports an association between certain patterns of drinking and the incidence of violence. However, even though harmful drinking and violent behavior may co-occur, there is no evidence showing that it is drinking that *causes* violence.

4.1.2 Causation

A causal relationship exists when one event (cause) is necessary for a second event (effect) to occur. The order in which the two occur is also critical.

- For example, for intoxication to occur, there must be heavy drinking, which precedes intoxication.

Determining cause and effect is an important function of evaluation, but it is also a major challenge. Causation can be complex:

- Some causes may be necessary for an effect to be observed, but may not be sufficient; other factors may also be needed.

---

3   See: http://www.um.dk/en/menu/DevelopmentPolicy/Evaluations/Guidelines/

- Or, while one cause may result in a particular outcome, other causes may have the same effect.

Being able to correctly attribute causation is critical, particularly when conducting an evaluation and interpreting the findings.

4.1.3 Confounding

To rule out that a relationship between two events has been distorted by other, external factors, it is necessary to control for confounding. Confounding factors may actually be the reason we see particular outcomes, which but may have nothing to do with what is being measured.

To rule out confounding, additional information must be gathered and analyzed. This includes any information that can possibly influence outcomes.

When evaluating the impact of a prevention program on a particular behavior, we must know whether the program may have coincided with any of the following:

- Other concurrent prevention initiatives and campaigns;
- New legislation or regulations in relevant areas;
- Relevant changes in law enforcement.
  - For example, when mounting a campaign against alcohol-impaired driving, it is important to know whether other interventions aimed at road traffic safety are being undertaken at the same time. Similarly, if the campaign coincides with tighter regulations around BAC limits and with increased enforcement and roadside testing by police, it would be difficult to say whether any drop in the rate of drunk-driving crashes was attributable to the campaign or to these other measures.

Addressing possible confounders is an important element for proper interpretation of results.

- However, it is often impossible to rule out entirely the influence of confounders.
- Care must be taken not to misinterpret the results of an evaluation and to avoid exaggerated or unwarranted claims of effectiveness. This will inevitably lead to loss of credibility.
- Any potential confounders should be openly acknowledged in the analysis of the evaluation results.
- It is important to state all results in a clear and unambiguous way so that they are easy to interpret.

4.2 Short- and Long-term Outcomes

The outcomes resulting form an intervention may be seen in a number of different areas, including changes in skills, attitudes, knowledge, or behaviors.

- Outcomes require time to develop. As a result, while some are likely to become apparent in the short term, immediately following an intervention, others may not be obvious until time has passed.
- It is often of interest to see whether short-term outcomes will continue to persist over the medium- and long-term.

> **KEY TERMS:**
>
> - **Short-term outcomes**
>
>   Most likely include changes in skills, attitudes, and knowledge.
>
> - **Medium-term outcomes**
>
>   Include changes in behavior and decision-making.
>
> - **Long-term outcomes**
>
>   Persistence of behaviors and broader lifestyle changes.

Evaluators should try to address short-, medium-, and long-term outcomes of an intervention separately.

- If the design of a program allows, it is desirable to be able to monitor whether its impact is sustained beyond the short term.
- Care should be taken to apply an intervention over a sufficiently long period of time so that outcomes (and impact) can be observed and measured.

Short- and long-term outcomes can be measured by using different methodologies for collecting data.

- **Cross-sectional** studies involve measurement at a single point in time after the intervention has been applied and allow short-term results to be measured
- **Longitudinal** study designs, on the other hand, follow progress over longer periods and allow measurements to be taken at two or more different points in time. They can help assess outcomes into the medium- and long-term

Unfortunately, the reality is that, for most projects, resources and time frames available are likely to allow only for the measurement of short- and perhaps medium-term outcomes.

4.3 Providing the Proper Context

Interpreting results is only possible in the proper context. This includes knowing what outcomes one can reasonably expect from implementing a particular intervention based on similar interventions that have been conducted previously.

For instance, when setting up a server training program, it is useful to know that such interventions have in the past helped reduce the incidence of violence in bars.

Therefore, once the intervention is over, if the results are at odds with what others have observed, it is likely that the program was not implemented correctly or that some other problem has occurred.

5. Reporting and Dissemination

How results of an evaluation are reported depends on the purpose of the report:

Is it to be used as a basis for repeating and implementing the intervention elsewhere? Is it to justify funding? Or is it to demonstrate that the intervention has worked (or has not worked)?

However, *any* comprehensive evaluation report must be clear, accurate, and easily accessible to the end user and should include the following:

- An **executive summary** presenting the main findings of the evaluation
- A **clear description of the intervention** being evaluated
- **Statement of purpose** of the evaluation and what was being measured (e.g., awareness, behavior change)
- A clear **explanation of the methodology** used, including data collection methods
- **Findings**, usually linked to particular program objectives against which performance is assessed (attention should be paid here to association, causation, and possible confounders)
- **Conclusions**, lessons learned, and recommendations
- **Annexes**, including any background information on the program or evaluation that may be of interest (e.g., the terms of reference and lists of people interviewed and documents reviewed)

How results are to be disseminated will also help inform the presentation of the results:

- It is important to decide on the number and type of outputs expected from the evaluation (e.g., report, summary brochures).
- More than one format may be required, depending on the composition of the target audience and key stakeholders (e.g., a comprehensive report for program funders and a short brochure to raise awareness of the activities among program target beneficiaries or others).

EVALUATION CHECKLIST: Disseminating Results

A number of dissemination methods may be needed to reach all evaluation target audiences. Possible approaches include:

- Sharing findings at meetings with reference groups of stakeholders
- Distributing evaluation results (e.g., final report, brochures) through conference presentations and information sharing at events
- Highlighting key findings in newsletters or other publications
- Publishing results in scientific journal articles
- Posting links to evaluation results on relevant websites
- Using social networking sites for updates and dissemination of information and results

Regardless of the dissemination strategy, the following are some simple and useful ways to present and report data.

Quantitative findings:

- Numerical data (e.g., percentages and rankings) are best presented as tables.
- Tables provide an easy overview of the results, and differences and similarities become quickly apparent.
- Where "yes" and "no" answers were used to measure outcomes, these should be grouped and added up so that a total number can be presented.
- Where respondents were asked to provide rankings, reporting is best done by calculating an average or mean value of the answers to each question. It is also useful to indicate how many people gave a particular answer.
- Where change is monitored over time, percentages can be used to show increases and decreases.
- If the sample size is large enough, additional statistical calculations can be done, for example, standard deviations from the mean or confidence intervals. These give an idea of how much variation there may have been among the answers and how much they may deviate from the average.

Qualitative findings:

Where responses were not given as numbers or values, other ways must be found to identify common themes and groupings.

- For example, results of focus groups can be reported as quotes, and consolidated into groups of similar responses or categories.

- It is important to identify patterns among the responses: Did respondents have similar views on different things? Did they have the same reactions and/or concerns?

- Grouping multiple responses and identifying themes help with interpreting the findings and understanding similarities and differences.

6. Evaluation Caveats and Pitfalls

Although evaluations are essential in determining whether a program has really worked, they are not always welcomed for several reasons:

- **evaluation carries risks**
- **results may not be transferrable**
- **evaluation may compete for tight resources**

6.1 Evaluation Carries Risks

Although this is not a desirable result, an evaluation may show that a program did not have its intended effect. It may show unexpected outcomes or impact that may not be as positive as expected.

However, unanticipated outcomes may actually be useful and should be included and reported. Not only is open and honest reporting of all results scientifically ethical and in the spirit of transparency, but unexpected outcomes may contain important lessons:

- They can be an indication that the assumptions upon which the program was based were wrong

- They can also be a sign that the program may not have been delivered properly and that there was a flaw in the design

- Unexpected results can also indicate that the data were not collected in the proper way

It is, therefore, important to look carefully at any negative or unexpected outcomes and identify what may have gone wrong. This will help with future design in the event that the program is repeated.

6.2 Results May Not Be Transferrable

Not all approaches to prevention are easily transferrable across cultures and to different settings.

- It is important to understand what is likely to resonate with a particular audience and what may be understood to be culturally acceptable.

It is also important to anticipate which approaches are likely to work and which may be doomed to failure.

- For example, implementing an intervention that relies on printed material where there is a high rate of illiteracy is not likely to be successful.

- Similarly, an intervention that relies on abstinence from drinking in a culture where alcohol consumption is well integrated into daily life is not likely to resonate well with the target audience.

6.3 Evaluation May Compete for Tight Resources

Evaluations are labor-intensive and require an investment of human and financial resources. Therefore, they may interfere with program activities and may require a trade-off with the actual delivery of the program.

However, careful planning can reduce evaluation costs.

- Integrating evaluation into the design and execution of the program can be helpful. This positions the evaluation as an integral part of the process and delivery, rather than as an interference.

In the long run, however, long-term benefits of proper evaluation will outweigh any short-term costs.

Annex A. A Stepwise Guide to Evaluating Prevention Programs

This Annex provides specific examples of how evaluations may be conducted for different types of programs aimed at preventing alcohol misuse. They outline a number of key steps to take and suggest the types of data that may be gathered, the questions that may be asked, and other parameters that may be applied.

The recommendations in this Annex are intended to be indicative only. The most appropriate approach will need to be identified and applied individually in each instance. It will be determined by the specifics of the initiative being assessed, particular conditions that may exist, cultural and other considerations, and available resources.

The Stepwise Guide is intended to complement ICAP's 2010 toolkit *A Guide to Evaluating Prevention Programs*. These two documents offer more comprehensive guidance to designing and implementing evaluations. This Annex applies the principles of evaluation to two scenarios that mirror the types of initiatives likely to be implemented by industry stakeholders—producers, trade associations, and SAOs. The case studies are:

Case Study 1: Evaluating an Awareness Campaign around the Use of Designated Drivers

Case Study 2: Evaluating a School-based Alcohol Education Program (Adapted from *Evaluation Guidelines for Education Interventions to Promote Responsible Drinking amongst Young People in Seven European Countries* by EFRD, www.efrd.org)

Each Case Study outlines the following phases needed for evaluation:

● Phase 1: Planning and Inception
● Phase 2: Data Collection
● Phase 3: Reporting and Dissemination

Case Study 1: Evaluating an Awareness Campaign around the Use of Designated Drivers

Background

Objective: To raise awareness about and encourage the use of designated drivers.

Target Audience: Young people of legal drinking and driving age who are likely to go out and consume alcohol in entertainment areas and venues such as bars, cafés, and nightclubs.

Approach: Information and educational material, including pamphlets and other reading materials and tools developed and distributed in venues where young people are most likely to gather.

Evaluation

The following is an outline of the phases that are required for planning, conducting, and reporting on the evaluation. It suggests the questions to ask, identifies the most appropriate approaches, and provides the timeline for implementing different elements.

Phase 1: Planning and Inception

The Purpose of Evaluation

This evaluation may have different purposes. For example:

1. To evaluate the impact of campaign on any or all of the following:
   a. Awareness of designated driver concept
   b. Use of designated drivers
   c. Incidence of alcohol-related traffic crashes
2. To share the results of the evaluation with key external stakeholders as a demonstration of responsibility and commitment to reducing alcohol-related harm.
3. To provide funders with evidence of performance and impact.

It is important to decide what the evaluation is intended to achieve so that it can be crafted accordingly.

Identifying Stakeholders

The stakeholders are all those who may be involved in some way in the evaluation or be recipients of its results.

External stakeholders may include:

● The target audience for the campaign, i.e., young people

● Partners who may be involved in developing the promotional materials (e.g., police, health professionals, insurance companies)

● Retail sector—those who own, manage, and work in the venues and establishments where the campaign will be implemented

● Media

Internal stakeholders may include:

● Program funders

● Organization's Board of Directors or governance bodies

Evaluation Methodology

The evaluation methodology will require the measurement of outputs, outcomes, and impact[4], using specific questions. It should be identified and clearly defined before the campaign is implemented.

The methodology to be used will depend on the resources that are available. This will determine whether the evaluation can be carried out externally or internally, as well as the number of data and time points that can be included.

A more detailed overview of the methodology is offered in **Phase 2: Data Collection**.

Evaluation Plan

Evaluation should be taken into account at all key stages of the campaign:

1. Before the campaign is launched
2. During its implementation
3. After the campaign has been completed

Developing a concrete step-by-step evaluation plan will help guide the process. This plan must be finalized well before the campaign is launched and should define activities scheduled for each phase of the evaluation.

Table 1 in ICAP's 2010 toolkit *A Guide to Evaluating Prevention Programs* outlines some of the steps that can be included in the evaluation plan; Table 3 of the toolkit offers a sample template.

Phase 2: Data Collection

The methodology needed for evaluation requires measurement of output, outcome, and impact. The data need to be collected at three different points:

1. Before the campaign
2. During the campaign
3. After the campaign

This allows measurements and data to be compared so as to determine impact.

1. Before the campaign

Establishing a baseline

Data should be collected to establish existing levels of awareness and behaviors against which to measure any changes that may result from the campaign.

Some possible questions include:

● What is the general level of awareness around drinking and driving?

● What is the level of awareness among respondents around designated driver schemes?

● How many respondents have used / are likely to use designated drivers?

● How many respondents have themselves been designated drivers / are likely to be designated drivers?

There are various ways to gather these data:

● Surveys conducted in premises where campaign was carried out

● One-on-one interviews with target audience, also on premise

● Focus groups drawn from target audience

Information is also needed on the demographics of the respondents.

● What age groups are represented among the respondents?

● What is the gender composition?

● Information on drinking patterns of respondents.

● Information on those who drink and drive.

● Information on those who are familiar with or likely to use designated driver schemes.

---

4   Definitions of what constitutes an intervention's outputs, outcomes, and impact are provided in ICAP's 2010 toolkit *A Guide to Evaluating Prevention Programs*.

Addressing confounders

Additional information is required to rule out other influences, including:

- Other concurrent prevention initiatives and campaigns around designated driver schemes
- New legislation or regulations around drinking and driving
- Changes in enforcement (e.g., increased police vigilance, road blocks, fines)

This information is needed before the campaign is launched. It allows any adjustments to be made to the study design so that such external factors do not interfere with the results.

Impact on objective measures, such as road traffic incidents, can be assessed through desk research, as can the analysis of official law enforcement statistics on crashes, fatalities, and breath-testing results at road blocks.

Any potential confounders need to be acknowledged in the analysis of the evaluation results and their likely effects discussed.

2. During the campaign

Measuring output

The materials used and the number of respondents targeted should be carefully tracked while the campaign is in progress.

- How many brochures or other promotional materials were printed or otherwise produced?
- How many were given to the target audience?
- How many separate venues were involved in promotion?
- How many times was information distributed?

3. After the campaign

After the campaign is finished, a second round of information-gathering will help with assessment of outcomes (and impact). These data can be collected through:

- Surveys conducted in premises where the campaign was carried out
- One-on-one interviews with target audience(s), also on premise
- Focus groups drawn from target audience(s)

Ideally, the same respondents used to set the baseline should also be used for data collection after the campaign.

It is important to also have a **control group** against which to measure any changes in the target group. Both groups should have a similar demographic composition. The difference is that that the control group will not have been exposed to the campaign.

Measuring outcomes

- Were the materials read by their intended audience? Was there interest in them?
- Are target audience members aware of the campaign and the information contained in the materials?

Measuring impact

The most desirable results relate to the impact that the campaign may have had on attitudes and behaviors of the respondents. In this case, the impact also relates to any changes that may have occurred in road traffic crashes as a result of the campaign. The same measurements that are taken from the target group of the campaign should also be taken from the control group. For example:

- Did the campaign change awareness about designated driver schemes among respondents? Did awareness increase/decrease/stay the same?
- What was the level of awareness among the control group?
- Were changes different among different groups of respondents (e.g, differences by age, gender)?
- Did the campaign change intent to rely on designated drivers? Would respondents be more/less likely to use a designated driver as a result of having been exposed to the campaign? How does this compare with the control group?
- Did the campaign result in a change in behavior? Among whom?
  - Are members of target audience more likely to use designated drivers as a result?
  - Have they changed their use of designated drivers?
- Did the campaign have an impact on road traffic incidents (crashes and fatalities)? By how much?

  (This sort of measurement is only realistic if the sample size used is large enough to effect a demonstrable change in objective measures like road traffic crashes.)

If possible, a followup data point should also be included. This should be collected some time after the campaign is concluded to allow assessment of whether any changes due to the campaign have persisted. Followup is particularly important in determining any effect on the behavior of respondents.

Reliability and consistency

To increase the likelihood that data collected are reliable and consistent across different time points, it is important to pay attention to the following:

- All data should be collected in the same way, using the same instrument (e.g., questionnaire), at each time point.

- Where one-on-one interviews or focus groups are used, the same individuals should be included, to the extent possible, in each data collection round.

- Where this is not possible, or where broader surveys are used, efforts should be made to maintain consistency among respondents: i.e., the same target groups should be used, and their demographic compositions should be kept as consistent as possible.

Phase 3: Reporting and Dissemination

Analysis of Findings

The purpose of the analysis phase is to transform the data collected into credible evidence about the development of the campaign and its performance. This process usually includes the following steps:

- Organizing the data for analysis (i.e., data preparation)

- Describing the data (e.g., generating findings of facts)

- Interpreting the data (e.g., assessing the findings against the evaluation criteria)

Collating and analyzing qualitative data can be a challenge. The following guidelines may be helpful:

- Writing up interviews and focus group discussions regularly and quickly (e.g., at the end of each day)

- Conducting regular analysis, combining data into key points

- Involving the whole team during the analysis process

- Using examples to illustrate key findings

- Using tables, photos, and other visual means to present key findings

Final Report

A well structured report should present findings in a clear and interesting way. At a minimum, it should include the following:

- The executive summary presenting the main findings of the evaluation

- An overview of the purpose of the evaluation and how it was carried out

- A clear explanation of the methodology used, including data collection methods

- Findings, usually linked to particular objectives against which performance is assessed

- Conclusions, lessons learned, and recommendations

- Annexes including the Terms of Reference,[5] list of people interviewed, and list of documents reviewed

Dissemination Strategy

It is important to decide on the number and type of outputs expected from the onset of the evaluation (e.g., report and summary brochure). More than one format may be required depending on which audiences will be targeted with the information. For example, a comprehensive report may be suitable for donors but a more concise brochure to raise awareness of activities may be sufficient for target beneficiaries or others.

If the target audience is the scientific community and the evaluation is to be used to help build the evidence base around particular types of interventions, data need to be carefully collected and analyzed and presented so they will withstand scientific scrutiny. Additional statistical analysis may be required.

There are various possible audiences for sharing the results, and they will shape the dissemination strategy:

- Sharing findings in a meeting with a reference group of stakeholders

- Distributing report through conference presentations

- Highlighting the key findings in newsletters and other publications

- Publishing journal articles (each journal will have its own guidelines and requirements)

- Posting links to your report on relevant websites

- Targeted e-mails and posts on social networking sites

---

5   A brief overview of what can be included in the Terms of Reference is provided in ICAP's 2010 toolkit *A Guide to Evaluating Prevention Programs*.

Case Study 2: Evaluating a School-based Alcohol Education Program[6]

Background

Objective: To educate school-aged children and young people about alcohol and its effects on the body; to promote responsible attitudes towards alcohol; to reduce the age at which drinking is initiated.

Target Audience: Secondary school students.

Approach: Educational materials about alcohol and drinking patterns will be used. These may be printed materials or may be web-based learning resources. Working together with schools and teachers, these materials will be integrated into the school curriculum.

Evaluation

The following is an outline of the phases that are required for planning, conducting, and reporting on the evaluation. It suggests the questions to ask, identifies the most appropriate approaches, and provides the timeline for implementing different elements.

Phase 1: Planning and Inception

The Purpose of Evaluation

This evaluation may have different purposes. For example:

1. To evaluate the impact of the program on:
   - Knowledge and understanding of alcohol and its impact on the body
   - Attitudes around drinking
   - Drinking rates, levels, and patterns
   - Age when drinking begins

2. To share the results of the evaluation with key external stakeholders as a demonstration of responsibility and commitment to reducing alcohol-related harm.

3. To provide funders with evidence of performance and impact.

It is important to decide what the evaluation is intended to achieve so that it can be crafted accordingly.

Identifying Stakeholders

The stakeholders are all those who may be involved in some way in the evaluation or be recipients for its results.

External stakeholders may include:
- The target audience for the program, i.e., young people of secondary school age
- Teachers involved in delivering this educational intervention
- Parents of young people participating in the program
- Other partners involved in developing and administering the educational tools (e.g., school administrators, Departments of Education)
- Media

Internal stakeholders may include:
- Program funders and partners
- Organization's Board of Directors or governance bodies

Evaluation Methodology

The evaluation methodology will require the measurement of outputs, outcomes, and impact[7], using specific questions. It should be identified and clearly defined before the campaign is implemented.

The methodology to be used will depend on the resources that are available. This will determine whether the evaluation can be carried out externally or internally, as well as the number of data and time points that can be included.

A more detailed overview of the methodology is offered in **Phase 2: Data Collection**.

Evaluation Plan

Evaluation should be taken into account at all key stages of the program:

1. Before the program is launched
2. During its implementation
3. After the program has been completed

Developing a concrete step-by-step evaluation plan will help guide the process. This plan must be finalized well before the campaign is launched and should define activities scheduled for each phase of the evaluation.

---

6   Adapted from: European Forum for Responsible Drinking (EFRD). (in press). *Evaluation guidelines for education interventions to promote responsible drinking amongst young people in seven European countries.* Brussels: EFRD. Available: http://www.efrd.org/

7   Definitions of what constitutes an intervention's outputs, outcomes, and impact are provided in ICAP's 2010 toolkit *A Guide to Evaluating Prevention Programs*.

Table 1 in ICAP's 2010 toolkit *A Guide to Evaluating Prevention Programs* outlines some of the steps that can be included in the evaluation plan; Table 3 of the toolkit offers a sample template.

Phase 2: Data Collection[8]

The methodology needed for evaluation requires measurement of output, outcome, and impact. The data need to be collected at different points:

1. Before the program is launched
2. During the program period
3. After the program is completed

This allows measurements and data to be compared so as to determine impact.

1. Before the program

Given the nature of the study and that it involves young people, parents' consent is required. This should be addressed as a first step before the program is put in place.

Also, all materials should be shared with teachers, and final adjustments made based on their input.

Establishing a baseline

Data should be collected to establish existing levels of awareness and behaviors against which to measure any changes.

A full description of questions and sample questionnaires are included in the EFRD *Evaluation Guidelines for Education Interventions to Promote Responsible Drinking amongst Young People in Seven European Countries*. However, some possible questions include:

- What is the level of awareness around drinking and alcohol, generally?
- What is the level of awareness among respondents around harms and benefits?
- How many respondents have consumed alcohol? At what age? How often do they drink?
- What are their general attitudes towards drinking?

Further qualitative data about attitudes and awareness can be collected through focus groups. Sample questions suitable for focus groups are outlined in the EFRD Guidelines.

Additional information can be gathered by conducting focus groups with parents to assess their views on their children's attitudes and behaviors toward alcohol.

Addressing confounders

Additional information is required to rule out other influences, for example:

- Assess whether any other information and programs are in place to which respondents may be exposed during this period;
- Validate the questionnaires to be used by administering them to a subgroup of respondents. This will determine whether any adjustments are needed to the research design or the questions themselves.

This information is needed before the campaign is launched. It allows any adjustments to be made to the study design so that such external factors do not interfere with the results.

Any potential confounders need to be acknowledged in the analysis of the evaluation results.

2. During the program

Measuring output

The materials used and the number of respondents targeted should be carefully tracked while the campaign is in progress. Some data to include:

- How many young people were involved in the educational program?
- What are the ages and gender composition?
- Which materials did they use (e.g., print, web-based)?

3. After the program

After the campaign is finished, a second round of information-gathering will help with assessment of outcomes and impact. These data can be collected through:

- Surveys conducted in premises where the campaign was carried out
- One-on-one interviews with target audience(s), also on premise
- Focus groups drawn from among the young people or their parents

It is important to also have a second group to measure. These should be young people who have not been exposed to the program. They are the **control group** and should be asked the same questions and given the same surveys as the **experimental group**.

It is important that the demographics of the control group closely match those of the group among whom changes are being measured.

Measuring outcomes

- Were the materials used and understood? Was there interest in them?
- Were students aware of the topics covered, and did they understand them?

Measuring impact

After the program has ended, questionnaires are to be administered to **both the control and experimental groups**. This will determine whether there has been any measurable impact on awareness and / or behavior.

- Survey questions that address drinking levels, patterns, attitudes, and awareness should be administered to both. (Sample surveys are included in the EFRD Guidelines).
- Qualitative one-on-one interviews should be conducted with a sub-group of respondents.
- Additional focus groups with both parents and young people are also helpful.

If possible, a **followup data point** should also be included. This should be collected some time after the program is conducted to allow assessment of whether any changes due to the campaign have persisted. Followup is particularly important in determining any effect on the behavior of respondents.

The questions above should be repeated and with the same respondent groups. This will allow insight into the following:

- Have changes been sustained over time? For how long? How much of the change is sustained?

(This will require follow-up measurements at different time points, with the same questions asked at each.)

Reliability and consistency

To increase the likelihood that data collected are reliable and consistent across different time points, it is important to pay attention to the following:

- All data should be collected in the same way, using the same instrument (e.g., questionnaire), at each time point.
- Where one-on-one interviews or focus groups are used, the same individuals should be included, to the extent possible, in each data collection round.

- Where this is not possible, or where broader surveys are used, efforts should be made to maintain consistency among respondents: i.e., the same target groups should be used, and their demographic compositions should be kept as consistent as possible.

Phase 3: Reporting and Dissemination

Analysis of Findings

The purpose of the analysis phase is to transform the data collected into credible evidence about the development of the program and its performance. This process usually includes the following steps:

- Organizing the data for analysis (i.e., data preparation)
- Describing the data (e.g., generating findings of facts)
- Interpreting the data (e.g., assessing the findings against the evaluation criteria)

Collating and analyzing qualitative data can be a challenge. The following guidelines may be helpful:

- Writing up interviews and focus group discussions regularly and quickly (e.g., at the end of each day)
- Conducting regular analysis, combining data into key points
- Involving the whole team during the analysis process
- Using examples to illustrate key findings
- Using tables, photos, and other visual means to present key findings

Final Report

A well structured report should present findings in a clear and interesting way. At a minimum, it should include the following:

- The executive summary presenting the main findings of the evaluation
- An overview of the purpose of the evaluation and how it was carried out
- A clear explanation of the methodology used, including data collection methods
- Findings, usually linked to particular objectives against which performance is assessed
- Conclusions, lessons learned, and recommendations
- Annexes including the Terms of Reference,[9] list of people interviewed, and list of documents reviewed

---

9   A brief overview of what can be included in the Terms of Reference is provided in ICAP's 2010 toolkit *A Guide to Evaluating Prevention Programs*.

Dissemination Strategy

It is important to decide on the number and type of outputs expected from the onset of the evaluation (e.g., report and summary brochure). More than one format may be required depending on which audiences will be targeted with the information. For example, a comprehensive report may be suitable for donors but a more concise brochure to raise awareness of activities may be sufficient for target beneficiaries or others.

If the target audience is the scientific community and the evaluation is to be used to help build the evidence base around particular types of interventions, data need to be carefully collected and analyzed and presented so they will withstand scientific scrutiny. Additional statistical analysis may be required.

There are various possible audiences for sharing the results, and they will shape the dissemination strategy:

- Sharing findings in a meeting with a reference group of stakeholders
- Distributing report through conference presentations
- Highlighting the key findings in newsletters and other publications
- Publishing journal articles (each journal will have its own guidelines and requirements)
- Posting links to your report on relevant websites
- Targeted e-mails and posts on social networking sites

Annex B. Additional Resources

General Guides to Program Planning and Evaluation

Evaluation Trust. (n.d.). *Introduction to evaluation methods and tools for the voluntary sector.* Bristol: Author. Available: http://www.evaluationtrust.org/tools/introduction

Higgins, J. P. T., & Green, S. (2009, September). *Cochrane handbook for systematic reviews of interventions. Version 5.0.2.* Available: http://www.cochrane-handbook.org/

McNamara, C. (2009). *Field guide to nonprofit program design, marketing and evaluation.* Minneapolis, MN: Authenticity Consulting. Available: http://www.authenticityconsulting.com/pubs/PG_gdes/PG_pubs.htm

Muraskin, L.D. (1993). *Understanding evaluation: The way to better prevention programs.* Washington, DC: U.S. Department of Education. Available: http://www2.ed.gov/PDFDocs/handbook.pdf

Saltz, R.F., & DeJong, W. (2002). *Reducing alcohol problems on campus: A guide to planning and evaluation.* Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism. Available: http://www.collegedrinkingprevention.gov/media/FINALHandbook.pdf

Taylor-Powell, E., Steele, S., & Douglah, M. (1996).*Planning a program evaluation.* Madison, WI: University of Wisconsin—Extension.

Interpreting Findings

Nadeau, J., & Camp, S. (2006). *Interpreting research studies.* Guttmacher Institute "In Brief" Series, No. 2. Washington, DC: Guttmacher Institute. Available: http://www.guttmacher.org/pubs/2006/07/27/IB_Interpreting.pdf

Martinic, M., & Leigh, B. (2004). *Reasonable risk: Alcohol in perspective.* New York: Brunner Routledge. Available: http://www.icap.org/Publications/ICAPBookSeries

Evidence of Effective Programs

Substance Abuse and Mental Health Service (SAMHSA). (n.d.). *SAMHSA's National Registry of Evidence-based Programs and Practices:* http://nrepp.samhsa.gov/

Alcohol Education Programs (General)

Foxcroft, D., Ireland, D., Lowe, G., & Breen, R. (2002). *Primary prevention for alcohol misuse in young people.* Cochrane Database of Systematic Reviews, Issue 3. Available: http://www2.cochrane.org/reviews/en/ab003024.html

Alcohol Education Programs (Social Norms)

Moreira, M. T., Smith, L. A., & Foxcroft, D. (2009). *Social norms interventions to reduce alcohol misuse in university or college students.* Cochrane Database of Systematic Reviews 2009, Issue 3. http://www.cochrane.org/reviews/en/ab006748.html

Programs against Alcohol-impaired Driving

Ditter, S. M., Elder, R. W., Shults, R. A., Sleet, D. A., Compton, R., Nichols, J. L., et al. (2005). Effectiveness of designated driver programs for reducing alcohol-impaired driving. *American Journal of Preventive Medicine, 28*(5S), 280–287. Available: http://www.thecommunityguide.org/mvoi/mvoi-AJPM-evrev-d-driver.pdf

Global Road Safety Partnership (GRSP). (2007). How to evaluate the programme. In GRSP, *Drinking and driving: A road safety manual for decision-makers and practitioners* (pp. 119–136). Geneva: GRSP.

Shults, R. A., Elder, R. W., Sleet, D. A., Nichols, J. L., Alao, M. O., et al. (2001). Reviews of evidence regarding interventions to reduce alcohol-impaired driving. *American Journal of Preventive Medicine, 21*(Suppl. 4), 66–88.

Programs for Young People

Clark, D. B., Gordon, A. J., Ettaro, L. R., Owens, J. M., & Moss, H. B. (2010). Screening and brief intervention for underage drinkers. *Mayo Clinic Proceedings, 85*(4), 380–391. Available: http://www.mayoclinicproceedings.com/content/85/4/380.abstract

European Forum for Responsible Drinking (EFRD) (in press). *Evaluation guidelines for education interventions to promote responsible drinking amongst young people in seven European countries.* Brussels: EFRD. Available: http://www.efrd.org

Foxcroft, D., Ireland, D., Lowe, G., & Breen, R. (2002). *Primary prevention for alcohol misuse in young people.* Cochrane Database of Systematic Reviews 2002, Issue 3. Available: http://www2.cochrane.org/reviews/en/ab003024.html

Jones, L., James, M., Jefferson, T., Lushey, C., Morleo, M., Stokes, E., et al. (2007, April). *A review of the effectiveness and cost-effectiveness of interventions delivered in primary and secondary schools to prevent and/or reduce alcohol luse by young people under 18 years old.* London: National Institute for Health and Clinical Excellence. http://www.nice.org.uk/nicemedia/live/11666/34522/34522.pdf

Moreira, M. T., Smith, L. A., & Foxcroft, D. (2009). *Social norms interventions to reduce alcohol misuse in university or college students*. Cochrane Database of Systematic Reviews 2009, Issue 3. http://www.cochrane.org/reviews/en/ab006748.html

Spoth, R., Greenberg, M., & Turrisi, R. (2009). Overview of preventive interventions addressing underage drinking: State of the evidence and steps toward public health impact. Alcohol Research and Health, 32(1), 53–66. Available: http://pubs.niaaa.nih.gov/publications/arh321/53-66.htm

Saltz, R.F., & DeJong, W. (2002). *Reducing alcohol problems on campus: A guide to planning and evaluation*. Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism. Available: http://www.collegedrinkingprevention.gov/media/FINALHandbook.pdf

Responsible Hospitality Programs

Graham, K., Osgood, D. W., Zibrowski, E., Purcell, J., Gliksman, L., Leonard, K., et al. (2004). The effect of the Safer Bars programme on physical aggression in bars: Results of a randomized controlled trial. *Drug and Alcohol Review, 23*(1), 31–41. Available: http://www3.interscience.wiley.com/journal/122412973/abstract

Ker, K., & Chinnock, P. (2008). *Interventions in the alcohol server setting for preventing injuries*. Cochrane Database of Systematic Reviews 2008, Issue 3. Available: http://www2.cochrane.org/reviews/en/ab005244.html

Screening and Brief Interventions

Clark, D. B., Gordon, A. J., Ettaro, L. R., Owens, J. M., & Moss, H. B. (2010). Screening and brief intervention for underage drinkers. *Mayo Clinic Proceedings, 85*(4), 380–391. Available: http://www.mayoclinicproceedings.com/content/85/4/380.abstract

Kaner, E. F.S., Dickinson, H. O., Beyer, F. R., Campbell, F., Schlesinger, C., Heather, N., et al. (2007). *Effectiveness of brief alcohol interventions in primary care populations*. Cochrane Database of Systematic Reviews 2007, Issue 2. Available: http://www2.cochrane.org/reviews/en/ab004148.html

Substance Abuse and Mental Health Service (SAMHSA). (2008, April). Brief Alcohol Screening and Intervention for College Students (BASICS). In *SAMHSA's National Registry of Evidence-based Programs and Practices:* http://nrepp.samhsa.gov/ViewIntervention.aspx?id=124