

Clean Air Partnership Evaluating CAP Programs

November 2005

Report Prepared By:

Clean Air Partnership

CLEAN AIR PARTNERSHIP
75 Elizabeth Street
Toronto Ontario
M5G 1P4

Telephone 416.392.6672
Fax 416.338.0616

cap@cleanairpartnership.org
www.cleanairpartnership.org



Executive Director:

Eva Ligeti, Clean Air Partnership

Information compiled by:

Jennifer Penney, Director, Research, CAP

© Clean Air Partnership, 2006

For more information, contact:

Clean Air Partnership
75 Elizabeth Street
Toronto, Ontario M5G 1P4
Canada
Tel: 416-392-6672
Email: cap@cleanairpartnership.org

Suggested citation:

Penney, J. 2006. Evaluating CAP Programs. Clean Air Partnership. Toronto, Ontario.

Evaluating CAP Programs

“One of the amazing things that has happened to evaluation is that it has pervaded the program world. Just about every organization that funds, runs, or develops programs now calls for evaluation. This is true locally, nationally, and internationally; it is almost as true of foundations and voluntary organizations as it is of government agencies. The press for evaluation apparently arises from the current demand for accountability. Programs are increasingly called on to justify their existence, their expenditure of funds, and their achievement of objectives. Behind the calls for accountability is an awareness of the gap between almost unlimited social need and limited resources.”

Carol Hirschon Weiss¹

1. INTRODUCTION

CAP has reached an important stage in its development. The organization is five years old and has an expanding and talented staff. CAP has a well-established presence in the Greater Toronto Area and good relations with all the governments operating in the GTA as well as with key environmental organizations.

CAP has several well-established education and social marketing programs, and has recently embarked on several new initiatives with the goal of facilitating collective, institutional and individual action to clean the air and halt climate change.

It is a good time for CAP to step back and assess the extent to which it has achieved its goals, and to analyse the strengths and weaknesses of its programs. We are also encouraged to undertake this evaluation by the Toronto Atmospheric Fund, a core funder of CAP, which is developing new evaluation measures for all the programs it supports.

2. WHY EVALUATE?

More and more environmental non-profits and other non-governmental organizations have begun to formally evaluate their programs and activities. This is partly because evaluation, like “business” planning, is being forced on NGOs by demands from funders to demonstrate results, and because of increased competition for funding. However NGOs are also coming to understand evaluation as an important part of effective program management. There are a number of reasons for this.

- a. Good evaluation processes can **promote clear thinking about the goals, design and implementation of programs**. When evaluation is included in the planning of programs or projects, it forces program designers to be more explicit about their goals and targets, describe how the planned activities can achieve these objectives and think about how these achievements will be measured.
- b. Outcome or impact evaluations help to **assess whether programs are achieving their intended purpose**. If the goal of a program is to reduce air pollution and greenhouse gas emissions by changing the behaviour of individuals and institutions, then the evaluation should try to assess how much behaviour has changed in order to gauge the success of the program.

¹ Weiss (2004), pages 1-2.

- c. Performance evaluations can also **identify the strengths and weaknesses of program design or implementation**. (Performance or implementation evaluation is more focused on how a program is carried out than on outcomes.)
- d. Participating in program evaluation can **help program staff to think critically about program design**. The efforts of staff to identify the strengths and weaknesses of program design or implementation can be fed back into the planning process to **encourage continuous improvements in program planning and implementation**.
- e. Evaluation may be designed to **measure the cost-effectiveness of programs**. A program may achieve the goals set for it, but a cost-effectiveness assessment may show that resources would be better invested elsewhere.
- f. Evaluation studies and performance monitoring may be required to **demonstrate accountability to funders** for their grants to the program. More and more funders are requiring outcome evaluation and some have begun to advise grantees about evaluation methods.²
- g. Evaluation can also **provide necessary information to Boards of Directors for informed decision-making**.
- h. And finally, evaluation can **strengthen the public image and credibility of an organization and its programs**.

3. CAUTIONS

Experienced program evaluators suggest a number of cautions for agencies attempting to develop outcome evaluations. These include the following:

- a. **Evaluation should not be used as an instrument to reward or punish**. Evaluation processes that are used to reward or to find fault may provide an incentive to “lie, cheat and distort”³ and lead to attitudes of defensiveness and justification rather than learning. Evaluation should be designed to promote learning and the improvement of programs. This requires an atmosphere in which managers and staff are encouraged to question what is being done and how it can be done better.
- b. **Good evaluation systems take time to develop and refine**. A recent guide to developing evaluation systems suggests that robust systems can take several reporting cycles to develop.

*“This does **not** mean that useful data cannot be collected and used fairly quickly. Rather it demonstrates the need to start as soon as possible, have tolerance for the necessary learning curve, and take steps to ensure that everyone is onboard for the journey.*

² The W.K. Kellogg Foundation, Hewlett Foundation, and United Way of America have all published guidelines for evaluating programs.

³ Perrin (2002), page 2

*Even if you could build the 'perfect system' on the first try, changes in the program environment, program goals, or technology would necessitate revision."*⁴

- c. **Evaluation is more difficult to do if it is not built into program design at the beginning.** Evaluation often requires the collection and monitoring of data that is much easier if it is planned at the beginning of a program.
- d. **Evaluation processes can be burdensome and can divert an organization's resources from valuable programs.** "Money that otherwise could go to program improvement ... instead needs to be devoted towards documentation, record keeping, report writing ...and related activities."⁵ Organizations undertaking evaluation for the first time will have to find a balance between the learning which evaluation can provide and the cost in time and resources that it will require, between the need for technically sound evaluation methodologies and the staffing, funding and workload realities that constrain nearly all environmental organizations.⁶

4. NEW APPROACHES TO EVALUATION

Program evaluation has historically been elaborated in an "accountability" framework focused on the proper use of funds, appropriate conduct of activities and the quantity and/or quality of program outputs or deliverables (materials produced and distributed, websites developed, research completed, events held, participants involved, etc.)

Assessments of this kind – sometimes called performance evaluation – are an important part of evaluation and will be described in a later section of this report. However, they do not usually answer the larger question of whether the ultimate goals of the program have been achieved.

An alternate approach – outcome evaluation – has gained considerable attention in the last decade. Outcome evaluation is focused on whether a program has achieved or is making progress towards achieving its goals.

As evaluators and funders have shifted to focus on outcomes, they have realized that organizations often make a large number of implicit and sometimes invalid assumptions about the link between their activities – however well conducted – and their goals.⁷ Often enough, their goals are not explicit or well defined.

As a consequence, many evaluation guides now advise organizations to clearly articulate:

⁴ SKG Consulting (2003), page 5

⁵ Perrin (2002), page 3

⁶ Albert et al (2005), page 6-3

⁷ A guide recently published by the Hewlett Foundation says: "Although intuitions are often the only realistic starting point for programs designed to bring about social change, many intuitively obvious interventions have turned out to be useless – even counterproductive. And unless it is implementing a strategy that has been well tested in similar contexts, the organization and its funders cannot know whether the program is making a difference or just spinning its wheels. Thus, we are constantly on the lookout for evidence bearing on the validity of the theories underlying our work..." Brest (2004), page xx

- The *goals* or expected *outcomes* of their programs;
- A *program theory* or *theory of change* which describes the pathways or connections between the activities of a program and the desired outcomes;
- A *logic model* which shows how the resources, activities and outputs of the organization will be mobilized to implement the theory of change;
- A description of the *measures* and *instruments* the organization will use to track progress as it implements the logic model and how it will evaluate success in achieving its goals.⁸

Each of these elements of program and evaluation planning is described in more detail below.

4.1 Articulating Goals or Outcomes

“Keep your eyes on the prize.”
Traditional Folk Song

Outcome evaluation examines how much progress a program or organization has made in achieving its objectives. This forces organizations to clearly identify short- and long-term goals. Organizations often confuse *outputs* or deliverables (such as reports, websites, workshops, conferences or policies) with *objectives* or *outcomes* (such as cleaner air or arresting human-induced climate change).

So an important initial step of an organization that undertakes outcome evaluation is to clearly identify goals and to keep them in mind in program planning. Many program planners and outcome evaluators suggest that both the design of a program *and* the design of an evaluation should start by defining the ultimate or long-term goals (often set out in the mission statement of the organization) and then work back to the intermediate and shorter-term outcomes that contribute to achieving the end goals of the organization.

For instance, achieving clean air, halting climate change and eliminating the health effects of air pollution and climate change may be described as CAP’s ultimate goals. Intermediate outcomes of specific programs, which are necessary for achieving the final outcomes, might include one or more of the following: reduced fossil fuel use; reduced emissions from transportation; reduced energy use in buildings; and a pronounced shift from polluting to non-polluting, renewable energy sources. Shorter-term outcomes of individual programs might be: increased use of bicycles and public transportation; preferential purchasing of Energy Star equipment; reduced idling of vehicles; improved energy efficiency of buildings; reduced exposure to air pollution or extreme heat and so on. Once the short- and long-term outcomes are clearly identified, then it is easier to think through how the outputs of a program might contribute to achieving them, and also easier to identify potential measures of program success.

4.2 Developing a Program Theory or Theory of Change⁹

Until recently, relatively few environmental or social programs had developed explicit theories that discuss the problems they are trying to address and set out the linkages

⁸ Adapted from Brest (2004).

⁹ Program theories are also sometimes called “causal models.”

between program activities and the changes necessary to solve the problem and contribute to long-term goals. Very often, programs have many unexplained steps (that some observers refer to as a “black box” or the “miracle in the middle”) along the path from program activities to outcomes. Because the pathway is not made explicit, flaws in the reasoning behind the program may be overlooked.

By contrast, program theories make explicit the ways in which the inputs and activities of a program are expected to achieve the outcomes. Program theories provide a logical framework for analyzing a problem, investigating interventions which could help solve the problem, thinking through the steps for developing a program, identifying expected outcomes and indicators that help assess the outcomes.

There are several approaches to elaborating a program theory or theory of change. This paper will discuss two possible approaches. One of these is based on social science theory. The version developed below draws on an approach recommended by the Kellogg Foundation (2004) and one used by the New York State Research and Development Authority (NYSERDA) for the planning and evaluation of their New York Energy Smart Program.¹⁰ The second is Carol Weiss’s “theory of change” approach.

4.2.1 Social Science Based Program Theory

“Successful programs create change and are built on a solid knowledge of what works – your program’s theory.”

W.K. Kellogg Foundation

A social science based program theory is developed as the result of a systematic investigation of information from multiple sources including: research analyzing the problem that the program intends to address, evaluation studies of comparable programs, the experience of program managers and staff, and existing theories in the social sciences.

Many funders require some sort of program theory from grantseekers, although they don’t usually use the term. A program theory is a description of the change strategy that your program supports, and usually includes:

- A description of the issue or problem the program addresses;
- Factors thought to be causing the problem;
- Which of these factors are addressed by the program;
- Target audience for the program and their status at the outset;¹¹
- Proposed strategy for intervening;
- Rationale about how the strategy should work, based on best practice research;
- Statement of assumptions about the selected strategy and why they will work in the proposed program setting;
- Sequence of hoped-for outcomes expected to result from the strategy; and
- Measurable indicators and explicit researchable (evaluation) issues.

NYSERDA’s program theory also incorporates an analysis of the driving and restraining forces that form the context of the program. (For example, rising energy costs may be a

¹⁰ See Albert et al (2005).

¹¹ This information may form important baseline data for an evaluation, and so is particularly valuable if it can be quantified.

driving force in the uptake of energy efficiency programs, while higher capital costs for energy-efficient equipment militate against energy conservation.) “Thoughtful consideration of these factors can help to identify risks, unknowns, and potential weaknesses within the program theory” that should be taken into account in program design.¹²

NYSERDA also conducts a “portfolio-level analysis” that shows how the different programs of the organization are expected to contribute to overall common goals. The agency produces diagrams that visually map out the links and pathways of the programs to achieving these goals.¹³

4.2.2 Theory of Change

“Theories represent the stories that people tell about how problems arise and how they can be solved.”

Carol Hirschon Weiss

The “theory of change” approach is more intuitive and less formal than the social science approach, and focuses more on the behaviour changes that a program tries to induce in its audience or clientele. Basically, the theory of change is a description of a logical sequence of steps linking a problem, the program designed to tackle the problem or some aspect of it, through to the changes it expects to result as part of the solution to the problem. This approach has been used by a number of American social service organizations in the planning and evaluation of programs for poor neighbourhoods and communities.

An illustration of the theory of change approach is set out below, as it might be applied to CAP’s Clean Air Consumer Guide.

- Air pollution is a problem we need to tackle.
- Use of fossil-fuel energy sources contributes to air pollution.
- The individual behaviours of consumers contribute to growing use of fossil-fuel energy sources and therefore to increased air pollution.
- People do not conserve energy because they are unaware of the air pollution impacts of energy use.
- They are also uninformed about how their behaviours contribute to unnecessary energy use.
- Education on the air pollution impacts will motivate people to reduce energy.
- Education on ways to reduce energy use will help them change the behaviours that increase energy consumption.
- Distributing the Clean Air Consumer Guide in a daily newspaper will reach a large number of people.
- A significant number of people who receive the Consumer Guide will read it.
- The information in the Consumer Guide is written in such a way that it persuades people who read it that air pollution is a problem they should be concerned about, and that it is strongly linked to their energy use.

¹² Albert et al (2005), page 6-3.

¹³ Other organizations – including Sheltair in its mapping of potential climate change adaptation programs for the City of Vancouver – have also used this kind of higher-level program theory to demonstrate visually how several programs can operate together to achieve common goals.

- The advice that the Consumer Guide provides for conserving energy is accurate and will, if followed, result in substantive energy use reductions.
- Few serious barriers exist to these changes in behaviour.
- These changes of behaviour will lead to an overall reduction in energy use and reduced emissions from fossil fuel sources of energy.
- Over the long term, this will lead to our ultimate goal – cleaner air.¹⁴

Some of the statements made above are pretty well-established and can be treated as fact rather than hypothesis (e.g. individual behaviours of consumers contribute to the growing use of fossil fuels and therefore to increased air pollution). However, there are also a number of problematic assumptions or leaps of faith that become evident when the theory of change for this program is made explicit (e.g. people do not conserve energy because they are unaware of the air pollution impacts of energy use, or that information on air pollution impacts and on ways to reduce energy consumption will lead to changes in energy-consuming behaviour).

When program designers elaborate an explicit theory of change and identify debatable assumptions, they may be encouraged to seek further evidence that would either support their assumptions or suggest more plausible ones (that may lead to alternative programs or program activities). This evidence might be sought from the experience of other practitioners who have designed and implemented successful programs, from the evaluation literature or elsewhere.

Using a theory of change to identify questionable assumptions may also help target specific issues for evaluation. In the case of the Clean Air Consumer Guide, evaluation might well focus on whether providing information to consumers changes their energy consumption behaviours, or whether the information is presented in a way that is effective in initiating behaviour change. If a preliminary evaluation demonstrates that behaviour change is occurring, then evaluators might further investigate the extent to which energy use and emissions might be reduced by the program. This information could provide a basis for deciding if the program is worth continuing, or whether resources should go to other, more promising programs.

Theory of change practitioners also argue that program developers who participate in a process of elaborating a program theory or theory of change “find this exercise as valuable a contribution to their thinking as the results of the actual evaluation.”¹⁵

4.3 Program Logic or Logic Models

If a program theory provides an overview of how the program is expected to create change, the logic model fills in the detail. A logic model requires a program planner to systematically analyze and describe the resources or inputs used in a program, the activities to be undertaken, the reach of the program (the audience or clientele), the outputs or deliverables, and the outcomes that are expected or hoped for.

¹⁴ Note: This theory of change is not intended to be definitive, but as an illustration of how a theory of change might be elaborated for a CAP program. A theory of change elaborated by CAP staff actually involved in the development, distribution and evaluation of the Clean Air Consumer Guide might look somewhat different. In fact, an evaluation process that involves several staff would likely produce somewhat varying ideas of a theory of change – and even identify points at which different staff members might be working at cross purposes because they have different implicit theories of change.

¹⁵ Weiss (1995), page 3.

A logic model may be shown graphically as a table, as in the following:¹⁶

HOW? (Resources & Activities)			WHO?	WHY? (Results)		
Inputs or Resources	Activities	Outputs or Deliverables	Reach	Direct or Immediate Outcomes	Intermediate Outcomes	Long-term Outcomes or Impacts
<p>The resources we use for the program:</p> <ul style="list-style-type: none"> ▪ Finances ▪ Staff ▪ Supplies ▪ Equipment ▪ Meeting venues ▪ Consultants ▪ Etc. 	<p>Things we do to create change:</p> <ul style="list-style-type: none"> ▪ Research ▪ Organize events & educational programs ▪ Produce written materials ▪ Create & support networks ▪ Etc. 	<p>Things we produce as a result of our activities:</p> <ul style="list-style-type: none"> ▪ Reports ▪ Publications ▪ Workshops ▪ Conferences ▪ Summits of decision-makers ▪ Working groups ▪ Advisory bodies ▪ Etc. <p>Note: Many expected outputs can & should be quantified at the start of a program, & monitored for evaluation purposes.</p>	<p>Groups or individuals who we:</p> <ul style="list-style-type: none"> ▪ Interact with ▪ Deliver our program to ▪ Seek to influence 	<p>The changes attitudes, knowledge and behaviour of those directly affected by the program, e.g.</p> <ul style="list-style-type: none"> ▪ Target audience knows more ▪ Audience acts on new information: <ul style="list-style-type: none"> ○ Adopts new policy ○ Takes steps to reduce energy use ○ Makes effort to reduce air pollutant emissions ○ Etc. 	<p>Often the shorter-term success criteria for the program, e.g.</p> <ul style="list-style-type: none"> ▪ Reduced energy use ▪ Reduced use of toxic products, processes &/or systems that contribute to air pollution ▪ Estimated or measured reductions in ghgs and/or air pollutant emissions ▪ Etc. 	<p>Ultimate program objectives:</p> <ul style="list-style-type: none"> ▪ Clean air ▪ Halt to human-induced climate change

While the model appears simple enough, it is often difficult for program planners using it for the first time. Most logic model guides suggest that, although the chart reads from left to right, it should be developed from right to left. That is, the desired outcomes should be developed first, then the outputs that would lead to the outcomes, the activities that are necessary to produce the outputs, and then the inputs required to undertake the activities.

A linear version of a logic chart can also be developed. The chart on the next page uses the proposed work on showcasing Toronto's green buildings as an example.

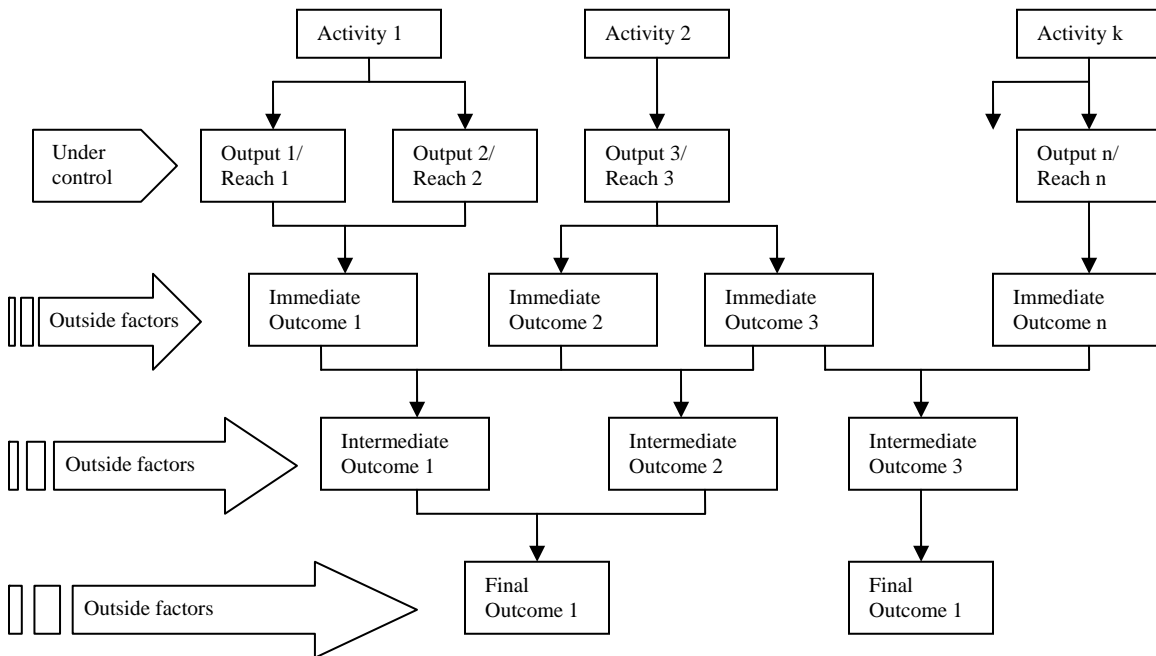
¹⁶ This format is adapted from McDonald (2002). The key elements of a theory of change could also be incorporated in the chart by adding a column to the left hand side of the table that spells out the main assumptions of the program, as the W.K. Kellogg Foundation did in the example in Appendix A.

4.3.1 Logic Chart for the Proposed Green Buildings Showcase Program

Inputs	<ul style="list-style-type: none"> ▪ Staff time ▪ Meeting costs ▪ Consultant costs for designing web site, map ▪ Printing and distribution costs for printed materials, etc.
Activities	<ul style="list-style-type: none"> ▪ Promote showcase idea ▪ Form a partnership with relevant organizations ▪ Seek agreement of Doors Open Toronto to create “Green Doors” event ▪ With partners, develop criteria for selecting buildings to showcase & identify candidate buildings ▪ Work with Doors Open Toronto & others to promote event ▪ Develop web site to showcase green buildings ▪ Gather information on showcase buildings, prepare case studies & put on website ▪ Design a map for locating showcase buildings & put on website ▪ Get agreement by showcase building owners to put up signage declaring the building a green or sustainable showcase ▪ Organize design competition for green buildings signage/identifier ▪ Research to estimate air pollution and ghg emissions reduction in showcase buildings
Outputs	<ul style="list-style-type: none"> ▪ Partnership of CAP, the City of Toronto, & organizations promoting sustainable building practices in the area ▪ Criteria for and inventory of green buildings in Toronto ▪ Green Doors Open event for Toronto with tens of thousands visiting green buildings ▪ Website with green building map, case studies, related CAP research & reports, & links to additional material and organizations ▪ Estimates of air pollution and ghg emissions reduction in green buildings
Reach	<ul style="list-style-type: none"> ▪ Partner organizations including City of Toronto, Canadian Urban Institute, Canada Green Building Council, Sustainable Buildings Canada, Toronto Society of Architects, Design Exchange, NAIMA, U of T School of Architecture, etc. ▪ Tens of thousands of City residents who participate in Doors Open Toronto ▪ Tens of thousands of people who visit the web site &/or read the printed materials, including building & design professionals, municipal staff, students & others
Immediate Outcomes	<ul style="list-style-type: none"> ▪ Strengthened partnerships with variety of organizations promoting sustainable building practices in the Toronto area ▪ Significantly increased awareness of and appreciation for existing green buildings in the area ▪ Increased understanding about the major design features of green buildings in this geographical area ▪ Increased understanding about the contribution of green building design to reducing air pollution & ghg emissions
Intermediate Outcomes	<ul style="list-style-type: none"> ▪ New collaborations with partners (e.g. development & promotion of green provincial building code) ▪ Increased popularity & demand for green building design & retrofits ▪ Increased support for green building standards in the City & Province
Long-term Outcomes	<ul style="list-style-type: none"> ▪ All new construction & major building retrofits incorporate green design ▪ Air pollution and ghg emissions from the building sector are substantially reduced

Where a program involves several activities, it can be mapped in a chart like the one below.¹⁷ This type of mapping can also be useful at getting a picture of work at a “portfolio level” as NYSERDA puts it, or of a whole organization. Note the increasing importance of outside factors as the outcomes go from the short- to the long-term.

4.3.2 Program Logic Chart



4.4 Indicators, Measures and Instruments to Track Progress and Evaluate Success

When outcomes, program theory, and program logic are clearly and explicitly developed, it becomes much easier to think about how to evaluate a program.

Evaluation questions can be formulated around four main links in the logic model:¹⁸

a. Did the inputs result in the planned activities being performed?

This question is straightforward to answer from internal data if the activities described in the program plan (logic model) are clearly delineated. The evaluation should include:

- An account of resources (financial, staffing and other, including the resources of program partners) actually required to mount the planned activities compared to those estimated during planning. Taking a good

¹⁷ Adopted from Mayne (1999), page 9.

¹⁸ These questions are adapted from List (2005).

look at actual versus expected inputs will be invaluable in developing staff capacity to effectively design and develop budgets for new programs.

- A simple report of planned activities that developed as expected. A detailed logic model developed during planning the program will make this report quite easy to produce.
- For activities that did not develop as expected, an examination of why not should be included in the evaluation. The reason may be as simple as an underestimate of the financial or human resources needed to mount a particular activity. Or that the promotion for an event was not targeted sufficiently. Or that essential partnerships could not be established for reasons which were not anticipated during the planning stages. It is very important to assess these problems in order to avoid making similar mistakes later.

b. To what extent did those activities produce the planned outputs and achieve the anticipated reach?

Part of the answer to this question can be answered from internal information. The answer will probably include quantitative as well as qualitative data, depending on the output. Quantitative data might include the number of educational events held; participants and/or municipalities taking part in summits, conferences or workshops; guidebooks distributed; hits on a website; etc. These numbers are easier to produce and more accurate if systems are set up at the beginning of a program to track and record the data. The quantitative assessment should be compared against numbers expected in the planning of the output.

For social marketing activities, the quantitative data will include message dissemination including direct distribution, distribution sites and amounts, distribution through intermediary organizations, media coverage and estimated audience size.

Qualitative assessment of significant outputs is also important, and this should involve a staff evaluation as well as an assessment by partners, stakeholders and/or audience.¹⁹ For staff, stakeholders and partners, a simple form could be used to collect an initial written response.²⁰ For important outputs, this can be followed up by a face-to-face discussion that will often elicit more insights than individual assessments.

Focus groups may also be a useful way to assess the effectiveness of outputs. (Focus groups can also be of considerable use at the activity level of the program work, to determine whether planned messaging, images and other potential outputs are effective.)

Audience assessments of the quality of events are usually fairly easy to arrange. Organizers can develop a simple survey to evaluate a conference or workshop,

¹⁹ To help avoid defensiveness on the part of program staff responsible for the output, positive feedback should be discussed first.

²⁰ This form could be as simple as a single page with three questions on it: What did you like about this (report/event/campaign/other output)? What didn't you like? What changes would you make?

distribute it to all participants, remind them to fill it out, and collect it as they leave. A short survey on a postage-paid card could be included in some publications. Surveys that are easy to answer are more likely to be completed. Evaluators usually recommend questions that can be answered using a five-point scale (e.g. very good, good, average, poor and very poor). A survey should always include the option for the respondent to provide additional comments. Sometimes the most valuable feedback comes from these comments. Surveys of the general public may require an incentive (e.g. a free compact fluorescent bulb) to improve response rates.

If an output – such as a social marketing campaign – is targeted at a particular audience, then survey questions should seek the demographic data that is necessary to determine if the message has reached its target audience. Information requests can be a measure of reach also, especially after a media campaign or a special distribution effort. A simple system could be set up to register information on the caller, what triggered their call and the type of information requested as a contribution to the assessment of public outreach.

Qualitative surveys of outputs may be used to produce quantitative data for an evaluation (For example: “70% of respondents found the guide to be very useful.” Or: “50% of respondents said they found no new information in the report.”)

The information sought on surveys will vary according to the output being evaluated, but it would be useful for the organization to develop a standard approach that can be incorporated in the planning of programs.

For a conference (or the Smog Summit) the evaluation may include assessment of the quality of the following elements:

- Event organization
- Venue
- Speakers
- Information and analysis
- Opportunities for discussion
- Opportunities for networking
- Planning for follow-up, etc.

For a research document, the survey might assess:

- Readability
- Information
- Thoroughness
- Analysis
- Relevance or usefulness of material

For a popular publication, the survey might collect information on:

- Attractiveness
- Readability
- How much of it was read
- New information
- Useful information

For a campaign, the survey might examine, among other things:

- Awareness of the campaign
- Understanding of the key message or messages
- Agreement with the message(s)
- Intention to change behaviour.

More collective forms of output evaluation may also be useful in some settings. For instance, CAP could conduct evaluations of key outputs – including the Smog Summit, key research and policy reports, or even the general quality of the meetings themselves – at some GTA-CAC meetings. An experienced evaluator could help in designing and facilitating these kinds of evaluation exercises.

Where partnerships or new working groups are part of the output of a program, they should also be evaluated. Evaluation issues might include: the extent of ongoing dialogue and information sharing; access to new audiences through the partnerships; provision of in-kind services; assistance in program development and/or implementation; coordinated messaging, etc.²¹

c. To what extent did those outputs result in the planned direct or immediate outcomes?

At this point we are crossing over from performance evaluation to outcome evaluation. The kinds of direct outcomes that CAP programs will typically be looking to produce and to evaluate include the following:

- Increased awareness of an air pollution or climate change problem (awareness change)
- Increased concern about air pollution or climate change problems (attitude change)
- Increased understanding of the factors that cause the problem (knowledge outcome)
- Increased understanding of the actions that could be taken to reduce the problem (knowledge outcome)
- Taking action to change practices that contribute to the problem (behaviour outcome)
- Promoting actions and policies intended to reduce the problem (behaviour outcome)
- Adopting policies designed to reduce the problem (behaviour outcome at the institutional or governmental level)

Most of these outcomes could be selected as indicators and used as measures of the success of a program. It can be pretty straightforward to design before-and-after surveys to assess changes in awareness, attitudes and knowledge about a problem as the result of a program, and this assessment is a legitimate part of outcomes evaluation. Unfortunately, changes in awareness, attitudes or knowledge are no guarantee that changes in behaviour that might reduce a problem will actually occur. And changes in behaviour – especially at the individual level – are typically more difficult to measure, in part because people tend to exaggerate behaviour changes that they know are desirable. This

²¹ Hersey et al (1999).

doesn't mean that surveys should not try to get an estimate of behaviour change, but that results should not necessarily be taken at face value. It might be wise to supplement a survey that indicates behaviour change as the positive output of a program, with other data that support this claim.

In the case of several anti-idling campaigns, telephone surveys were supplemented by observation of idling behaviour at active campaign sites. (In the research, observations demonstrated that there was no significant behaviour change among drivers that could be linked to the campaign. As a follow-up to this kind of result, evaluators and program staff should investigate why there was no significant behaviour change observed. It could be because the campaign was insufficiently resourced and simply didn't reach enough people, because weather conditions were different on the days when the before-and-after observations were conducted, or for other reasons. An evaluation that acknowledges failing to produce or to measure expected behaviour change, and identifies good reasons for the failure, allows for organizational learning that can improve subsequent programs.)

Changes of behaviour at an institutional level may be easier to measure, unless they are made up of actions of individual staff such as turning off computers at the end of the workday. If a CAP program develops model by-laws or policies for GTA municipalities, it is fairly easy to use interviews with the appropriate officials and to review publicly available documents to determine the extent to which these policies have been adopted. (For more complex policies, the adoption of policies may be regarded as an intermediate outcome, because of the longer time frame.)

It is also possible to evaluate very specific, localized behaviour change programs such as pilot projects, by other means. For instance, a program that promotes the use of compact fluorescent bulbs in a particular geographical area might be able to estimate its impact on behaviour by monitoring sales of the bulbs in the vicinity.

d. To what extent did those direct outcomes (if achieved) produce the desired intermediate or long-term outcomes?

Usually the short-term outcomes are the easiest to develop measures for, and to evaluate. The further out on the results chain, the more difficult it is to evaluate the impact of a program. This is because the intermediate and long-term results take more time to manifest and also because a greater range of external factors will inevitably influence longer-term results.

An evaluation of whether a program is achieving its longer-term goals may require the services of an external evaluator and will necessarily be more costly. It may be worthwhile for CAP to seek funding to undertake such an evaluation from time to time, especially for long-term programs or programs which represent a standard strategy of the organization. A credible (and positive) evaluation will also help in securing funding for ongoing programs.

If an external evaluator is used, it is important to involve staff in the design and where possible in the implementation of the evaluation, to maximize learning

about evaluation instruments and processes, and to increase organizational capacity.

A variety of instruments can be used to assess longer-term outcomes. Often a mix of instruments is necessary.

Modelling emissions reductions may be a useful tool in intermediate and longer-term outcome evaluation, as well as in setting quantitative targets in program planning. The behaviour change that a program achieves will form data inputs from which the modelling program will generate energy savings and reduced emissions estimates. However, its value will depend on a realistic estimate or measure of behaviour change resulting from the program.

Whether a program achieved its long-term outcomes is the most difficult – and most expensive – question to answer. Modelling data might be used (to estimate decreases in energy use, air pollutant or greenhouse gas emissions). Official statistics such as air pollution measures or greenhouse gas emission reductions, where available, might also be used in providing answers to this question.

5. Conclusions and Recommendations

The evaluation literature demonstrates that there are many benefits to evaluating social and environmental programs. It also makes clear that good program planning and evaluation are inextricably linked. The recommendations outlined below draw on this information.

a. That CAP adopt a standard program planning approach that includes:

- Clear identification of short-, intermediate and long-term outcomes for each program;
- Development of a program theory that identifies and provides a rationale for assumptions about how outcomes will be produced by the program;
- A logic model that lays out in detail the inputs, activities, outputs and outcomes for the program;
- Evaluation planning for all major projects, ensuring that data collection and tracking is planned in advance.²²

b. That CAP undertake capacity building for staff on program evaluation:

This might include workshops run by for staff and interested board members on evaluation strategies and instruments.

c. That CAP seek funding to undertake a formal external outcome evaluation of a key program or one which represents a standard approach for CAP.

This may be a social marketing program, the Smog Summit or other program in which CAP invests a great deal of resources. It might be useful to choose a program for which greenhouse gas emissions modelling could be incorporated in the evaluation.

²² Budget lines for tracking and evaluation should be included in most program budgets.