# TABLE OF CONTENTS

**NOTES** ...........................................................................................................1

**EXECUTIVE SUMMARY** ..................................................................................2
  Underlying Values and Principles .................................................................2
  Concepts .......................................................................................................2
  Approaches .................................................................................................2
  Benefits, Challenges, Suggestions For Success ........................................2
  Incorporating Evaluation into an Organization ............................................3

**INTRODUCTION** .......................................................................................4
  Purpose of this Review ................................................................................4
  Methods .......................................................................................................4
  Context .........................................................................................................5
  Contents .......................................................................................................5

**GUIDING PRINCIPLES** ..............................................................................7
  Values and Principles ................................................................................7
  Evaluation Standards ..................................................................................8
  Evaluator Ethics .........................................................................................9
  Research Ethics ..........................................................................................9
  Organizational Practice Guidelines ............................................................9

**EVALUATION CONCEPTS** ...................................................................11
  Defining Evaluation ....................................................................................11
  Evaluation Types .......................................................................................13
  Other Evaluation-Related Concepts ...........................................................19

**OVERVIEW OF APPROACHES** ............................................................23
  Steps to Guide Evaluation Practice ..........................................................23
  Summary of Major Evaluation Approaches ............................................26
  Results-Focused Evaluation .....................................................................28
  Utilization-Focused Evaluation .................................................................39
  Collaborative Evaluation .........................................................................40
  Balanced Scorecard ..................................................................................45
  Appreciative Inquiry .................................................................................49
  Other Approaches ....................................................................................54
  Summary Note on Evaluation Approaches ...............................................55

**EVALUATION IN PRACTICE** .................................................................56
  Benefits .......................................................................................................56
  Challenges ..................................................................................................56
  Suggestions for Success ............................................................................57
  Incorporating Evaluation into an Organization .........................................60

**REFERENCES** ..........................................................................................63
NOTES

- This review was prepared for the Saskatchewan Ministry of Education in 2008 to assist the development of a guiding framework for the Saskatchewan Ministry of Education’s evaluation activities.
- The Saskatchewan Ministry of Education holds the copyright to this document.
- Results were based on several key informant interviews and a literature review.
- The assistance of people in a number of ministries, in particular the Saskatchewan Ministry of Education, is gratefully acknowledged.
- Excerpts from the original document are included here with permission of the Saskatchewan Ministry of Education.
- Discussion of the specific Saskatchewan context and a possible Saskatchewan Ministry of Education framework are omitted from the current document.
EXECUTIVE SUMMARY

Underlying Values and Principles

Every framework is based on a set of underlying values and principles. Making values and principles explicit increases the likelihood they will be integrated into practice and promotes a reflective approach, resulting in greater understanding of, and ongoing revisions to, values and practice. Some documents translate values and principles into sets of guidelines, ethical principles or standards. For example, the Joint Committee on Standards for Educational Evaluation’s most recent draft of *Program Evaluation Standards* (2007) focuses on four categories: feasibility; propriety (ethics); accuracy; and utility (usefulness).

Concepts

Evaluation concepts are understood in very different ways, making it important to define them in order to ensure high quality communication. The concepts include evaluation itself, the various types of evaluation (ranging from process and outcome evaluations to meta-evaluation), and other evaluation-related concepts, such as: baseline, stakeholder, qualitative, quantitative.

Approaches

Many evaluation approaches exist. The literature suggests that no one approach is best for all situations as each approach has its own particular set of strengths and cautions. Often two or more approaches are combined when conducting an evaluation. Each approach has an associated set of steps designed to guide evaluation processes and activities; there is a large degree overlap in the suggested steps, although the nature of the methods and tasks related to each step vary. Major evaluation approaches include: results focused (goals based, goals free, theory based); utilization focused; collaborative; balanced scorecard; appreciative inquiry.

Benefits, Challenges, Suggestions For Success

Considering evaluation-related benefits, challenges and suggestions for success increases the likelihood of an evaluation framework successfully guiding an evaluation. Potential benefits range from increased capacity to improved outcomes. Challenges range from lack of resources to measurement issues. Examples of suggestions for success include planning for evaluation from the start of a project, reviewing processes as well as outcomes, and using a variety of methods and designs.
Incorporating Evaluation into an Organization

Examples of guidelines for incorporating evaluation into an organization include the following: take context into account; ensure adequate resources; foster an evaluation-friendly culture; integrate evaluation into ongoing processes and activities; define roles and responsibilities.
INTRODUCTION

Purpose of this Review
This review provides preliminary information to support the development of a guiding framework for Saskatchewan Ministry of Education evaluation processes and activities. This review uses the term “framework” to mean a set of steps to operationalize a particular evaluation model or approach. In this review the terms “approach” and “model” are used interchangeably.

Methods
Two methods were used to collect information for this review: a literature review and key informant interviews.

literature review
The following databases and resource websites were searched for potential sources of relevant information:

- Education Resources Information Centre (ERIC) database
- Web of Science database
- Google general
- Google Scholar
- websites of several provincial ministries of education: British Columbia, Alberta, Manitoba, Ontario and Quebec
- Government of Canada website
- numerous other websites, using a cascade approach (following one link leads to other links which lead to more links)

Where possible the key phrase “evaluation framework(s)” or “evaluation model(s)” was used. In some cases other key terms were added to the search such as “education” or “business.” The searches were restricted for the most part to the last 10 years (1998-2008) and English language documents.

Key informants suggested some documents; others were identified by reviewing the author’s personal collection.

The documents included in this review were selected to ensure information related to a variety of perspectives:

- business, human service and education sectors
- Saskatchewan government ministries
other provinces
other countries
overviews of evaluation approaches, models and frameworks
individual frameworks with defined steps

(See the Reference section for the complete list of documents referred to in this review.)

**key informant interviews**
Six key informants were interviewed, representing the Ministry of Education and other Saskatchewan ministries. The purpose of the interviews was to gain information about the use of evaluation frameworks in the provincial government, ideas about evaluation frameworks in general, and best ways to implement an evaluation framework in a ministry. Key informant statements included in this review were checked with key informants to ensure interview comments were interpreted accurately.

**limitations**
Due to short timelines an in-depth review was not possible. Some information may be incomplete. For the most part secondary sources were reviewed rather than original sources. (Secondary sources refer to material whose authors are not the creators of the content which they are discussing.)

**Context**
In many sectors program or policy evaluation is more the exception than the rule for a number of reasons, ranging from time and resource issues to a perception that evaluation is not worth the effort. At the same time, recognition is growing that evaluation is important to meet accountability requirements on the one hand and to improve policies and programs on the other.

That governments should apply themselves to continuously improving their performance, particularly with respect to the management of public funds and the stewardship of public assets, appears now to be a generally accepted principle of good governance. (Davies1999)

**Contents**
This review discusses the key components that compose a comprehensive evaluation framework:

- guiding principles
- concepts
- models/approaches
- steps for operationalizing the models or approaches
- understanding evaluation in practice: positives to encourage and negatives to avoid, challenges and solutions, and prerequisites for success

Only the basic evaluation model or approach and the operational steps are strictly necessary to a framework. However, each of the other components adds considerably to the understanding and appropriate use of a framework.
GUIDING PRINCIPLES

Every framework is based on a set of underlying values which translate into principles. Often these values and principles are implicit. The benefits of making them explicit are two-fold:

- Naming values and principles makes it easier to ensure that they are integrated into processes and activities.
- The process of identifying and defining values and principles promotes a reflective approach where underlying values and principles are examined. In some cases they are revised as a result of this reflection; in other cases a deeper understanding and commitment to them results. Further, processes and activities may change as a result of this reflection.

This section includes examples of different kinds of guiding principles.

Values and Principles

The guiding principles for evaluation in education adopted by this Committee emerged from the fundamental recognition that education is an interactive process. Schools and programs are guided by values aimed at shaping the academic, social, personal and career growth and development of students.

(Minister’s Advisory Committee on Evaluation and Monitoring 1989)

This section provides examples from two documents. The first document, *Saskatchewan Learning: Pre-K-12 Continuous Improvement Framework Guide* (Version Two 2007), contains a set of continuous improvement principles, for example:

Because we believe in equity and fairness we are committed to being sensitive and respectful to socio-demographic differences such as culture, gender, economic status and geographic location.

Because we believe in openness and honesty we are committed to open and transparent practices that honestly portray strengths and weaknesses in the system.

According to this document, “The principles that must guide this process of discovery, renewal and improvement are based on commonly held beliefs and practices that have guided the Pre-K-12 system through many initiatives in the past.”

The list below extracts the values on which this document’s principles are based:

- quality and excellence
- shared responsibility
openness and honesty  
equity and fairness  
comprehensiveness  
informed decision making  
authenticity and validity  
professionalism  
appropriate resources and staff development  
long-term planning for improved learner outcomes  
parental and community engagement

The second document, *Evaluation Policy and Framework* (Saskatchewan Post-Secondary Education and Skills Training 1998), provides an example of “values and beliefs about evaluation”:

- accountability  
- continuous improvement  
- program/service ownership, i.e. part of culture where evaluation is valued and integrated into operations  
- collaboration and inclusive effort, with stakeholder input sought and respected  
- adherence to evaluation standards, i.e. respect the rights of individuals and follow ethical and technical standards set out by CES (Canadian Evaluation Society)

**Evaluation Standards**

The Joint Committee on Standards for Educational Evaluation has a recent draft of *Program Evaluation Standards* (Draft 3 2007). Previous versions of this set of standards are highly respected and often cited.

The standards provide sound guidelines to follow when having to decide among evaluation options. In particular, the standards help avoid creating an imbalanced evaluation (e.g., one that is accurate and feasible but not useful or one that would be useful and accurate but is unethical and therefore infeasible). These standards can be applied both while planning an evaluation and throughout its implementation. (Milstein et al. 2000)

The standards are organized into the following four categories:

- **feasibility**: e.g. use resources efficiently and balance vested interests  
- **propriety (ethics)**: e.g. protect human and legal rights and be fiscally responsible  
- **accuracy**: e.g. use sound qualitative and quantitative methods and report truthfully  
- **utility (usefulness)**: e.g. align evaluation purposes with stakeholder and user needs and encourage participants to revise their understandings
Evaluator Ethics
The Canadian Evaluation Society’s *Guidelines For Ethical Conduct* focus on the evaluator’s conduct in the following three areas:

- competence
- integrity
- accountability

Research Ethics
The *Research Ethics Training Curriculum* (Rivera et al. undated) describes the “3 basic principles that are considered the foundation of all regulations or guidelines governing research ethics”:

- **Respect for persons**: “refers to the respect of the autonomy and self-determination of all human beings; acknowledging their dignity and freedom. An important component of this principle is the need to provide special protection to vulnerable persons” such as children or people living in poverty. “Respect for persons is embodied in the informed consent process. Informed consent is designed to empower the individual to make a voluntary informed decision regarding participation in the research. Potential research participants must fully comprehend all elements of the informed consent process.”

- **Beneficence**: “Beneficence makes the researcher responsible for the participant’s physical, mental and social well-being as related to the study...The risks to a person participating in a research study must be weighed against the potential benefit to the participant and the importance of the knowledge to be gained. In any case, all risks should be kept to a minimum.”

- **Justice**: “The principle of justice forbids placing one group of people at risk solely for the benefit of another. For instance, justice would not permit using vulnerable groups—such as minors, poor people, or prisoners—as research participants for the exclusive benefit of more privileged groups.”

Organizational Practice Guidelines
The *Suggested Guidelines* section of the IDM Manual (Kahan & Goodstadt 2005) provides guiding principles for a comprehensive range of organizational processes and activities, ranging from values and ethics to research/evaluation. For example, the evidence guidelines state that:

- **Sources of evidence**: should “derive from a wide variety of sources, including all key stakeholders and relevant key informants, and from a wide variety of methods”

- **The nature of evidence**: should “include the relationship between processes, activities and results/outcomes” (that is, not be restricted to identifying outcomes only)
The use of evidence: should “contribute to continuous learning and knowledge development, that is, lead to: a broad and complete picture of what is happening, insights into why things happen, understanding of how we might make things happen in a better way”

The document Business Process Improvement (BPI) And Program Review Framework (Saskatchewan Industry and Resources 2004) contains principles which “outline how P&E (Planning and Evaluation) and project participants will participate and manage projects under this framework”:

- Participation: encourage participation “by all who wish to participate and/or who might be affected by the review.”
- Decision Making: “Projects will utilize a structured decision-making process.”
- Value People: “Projects are not intended to result in a loss of employees but may result in employees being re-deployed to other activities within the department.”
- Measurement: for accountability; measures should be accurate, consistent, flexible, comprehensive but not onerous
- Integrated Program/Process Planning and Evaluation: incorporated into yearly business plans
- Ethical Conduct/Openness: consider ethical implications, respect and protect rights of participants
- Program/Process Focus: focus on improving program, activity or process
- Clear and Accurate Reporting of Facts and Review Results
- Timely Communication of Information and Review Results to Affected Parties
- Multi-Disciplinary Team Approach: include a range of knowledge and experience; seek assistance from outside of the team as required
- Customer and Stakeholder Involvement: “External and internal customers and stakeholders related to a project should be identified and consulted, if possible, throughout the project.”
EVALUATION CONCEPTS

Any framework designed to promote an understanding of evaluation in Saskatchewan should include a common conception of what evaluation is. (Minister’s Advisory Committee on Evaluation and Monitoring 1989)

Evaluation concepts are understood in very different ways. For this reason it is important to define them so everyone will attach the same meaning to them, or know what the differences are, thereby improving communication. An overview of common evaluation concepts follows. For an example of a complete set of evaluation definitions, see the Glossary of Key Terms in Evaluation and Results Based Management (Development Assistance Committee [DAC] Working Party on Aid Evaluation 2002).

Defining Evaluation

Many definitions of evaluation exist. Most definitions refer to program evaluation; some refer to project or policy evaluation. Some definitions use one term to refer to all types of evaluation, for example including policy and program evaluation under the umbrella title of policy evaluation. This section provides a number of examples of evaluation definitions.

Evaluation is often defined as an activity that judges worth. For example, evaluation is:

…the determination of merit, worth, or significance… (Scriven 2007)

A course of action used to assess the value or worth of a program. (Farell et al. 2002)

Some definitions include the notion of improvement. For example:

[Evaluation is] a set of research questions and methods geared to reviewing processes, activities and strategies for the purpose of improving them in order to achieve better results. (Kahan & Goodstadt 2005)

Policy evaluation is a family of research methods that are used to systematically investigate the effectiveness of policies, programmes, projects and other types of social intervention, with the aim of achieving improvement in the social, economic and everyday conditions of people’s lives. (Government Social Research Unit 2007)

The real purpose of an evaluation is not just to find out what happened, but to use the information to make the project better. (Community Tool Box undated)

Evaluation is an integral part of all aspects of the educational process and its major purpose is to improve instruction and student learning...Evaluation is the reflective link between the dream of what should be and the reality of what is.
This reflective link or activity will contribute to the vision of schools which are self-renewing, with mechanisms in place to support the ongoing examination and renewal of educational practices. Evaluation procedures must contribute to that vision. (Minister’s Advisory Committee on Evaluation and Monitoring 1989)

Other definitions focus on results. For example:

Project evaluation involves measuring and assessing the impact of the project based on specific criteria:
– effectiveness: the degree to which objectives have been attained, determined by comparing the results actually obtained to the results targeted
– relevance: the relationship between the objectives set and the needs to be met
– efficiency: the relationship between the quantity and quality of educational goods and services provided and the means used to obtain them: types of resources, organization and operation (Direction des services aux communautés culturelles and École montréalaise, Ministère de l'Éducation undated)

Evaluation is the process of assessing the impact of a project, programme or policy while it is in operation, or after it has come to an end. It involves consideration of the economy, efficiency and effectiveness of the project to determine whether the original objectives have been achieved. These will have been identified at the project initiation stage and documented in the business case. Evaluation brings to the fore the lessons to be learnt for the future which, in turn, should be fed into future decision-making. Evaluation does not seek to create blame for what did not go well. (Department of Health 2002)

Note that the last definition includes the notion of improvement in addition to results; some definitions include more than one focus. A definition of evaluation that incorporates most of the ideas mentioned in the above definitions follows:

The systematic and objective assessment of an on-going or completed project, programme or policy, its design, implementation and results. The aim is to determine the relevance and fulfillment of objectives, development efficiency, effectiveness, impact and sustainability. An evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision-making process of both recipients and donors. Evaluation also refers to the process of determining the worth or significance of an activity, policy or program. An assessment, as systematic and objective as possible, of a planned, on-going, or completed development intervention. Note: Evaluation in some instances involves the definition of appropriate standards, the examination of performance against those standards, an assessment of actual and expected results and the identification of relevant lessons. (Development Assistance Committee [DAC] Working Party on Aid Evaluation 2002)

The range of definitions can be confusing. However, Patton (2003) cuts through the confusion to make this point:
I feel no need for a definitive, authority-based definition [of evaluation], and I do not expect one to emerge despite the lamentations of those unsettled and confused by our diversity. It just means that we each bear the responsibility to be clear about what we mean by evaluation.

**Evaluation Types**

The literature lists a number of different types of evaluation. Definitions of these evaluation categories sometimes vary. The left-hand column in the table below lists key evaluation types as described in the document *Evaluating Crime Prevention through Social Development Projects: Handbook for Community Groups* (Public Safety and Emergency Preparedness Canada 2006). The right-hand column provides notes on the definition of the evaluation type based on information from other sources. The current trend in the literature is to advocate use of a range of evaluation types rather than any one type in particular.

**Table 1: Types of Evaluation**

(Note: All references in the table below are quoted verbatim.)

**Preliminary Evaluation**

**Handbook definition**

- A needs assessment is used to learn what the people or communities that you hope to reach might need in general or in relation to a specific issue. For example, you might want to find out about safety issues in your community, about access to services, or about the extent to which your community is dealing with a certain type of crime or a certain form of victimization.

- A resource assessment is used to assess the resources or skills that exist among the people or communities with which you hope to work. It is often conducted alongside a needs assessment. Resource assessments identify the skills that community members can

**Notes**

Other categorizations of evaluation types combine preliminary reviews or assessments into one general category. For example: Design evaluation is concerned with the pre-programme data collection such as baseline surveys, feasibility studies and assessment of pilot projects. (Walden & Baxter 2001)

Sometimes these preliminary tasks are designated as part of formative evaluation. For example:

Formative evaluation includes such things as
- needs assessments,
- evaluability assessment (analysis to determine if your program’s intended outcomes are able to be evaluated),
- program logic models,
- pre-testing program materials, and
- audience analysis. (Van Marris and King 2007)
Handbook definition  
contribute to a project and resources such as community space, in-kind and financial donations, volunteer time, and other attributes that can be tapped by your crime prevention project.

Notes  
• An evaluability assessment is done to determine whether a project is ready for a formal evaluation. It can suggest which evaluation approaches or methods would best suit the project.

Formative Evaluation  
Handbook definition  
• Also known as process evaluation, a formative evaluation tells how the project is operating, whether it is being implemented the way it was planned, and whether problems in implementation have emerged (for example, it might identify that a project is reaching a less at-risk group than it intended, that staff do not have the necessary training, that project locations are not accessible, or that project hours do not meet participant needs.).

Notes  
Other definitions restrict formative evaluations to a project’s planning or start-up phase. For example, formative evaluation:

is carried out during the early stages of the project before implementation has been completed. It focuses on ‘process’ issues such as decision-making surrounding the planning of the project, the development of the business case, the management of the procurement process, how the project was implemented, and progress towards achieving the project objectives.” (Department of Health 2002)

focusses on programs that are under development. It is used in the planning stages of a program to ensure the program is developed based on stakeholders needs and that programs are using effective and appropriate materials and procedures. (Van Marris and King 2007)

When formative evaluation is defined as applying to the project or program’s beginning phase, “process” or “implementation” evaluation is often used to refer to evaluations of processes or activities occurring in later phases of a project or program.
**Process Evaluation**

**Handbook definition**
- A process evaluation reviews project development and examines project activities to assess whether the project is being offered in the way it was intended and to identify areas where project administration and delivery can be improved.

**Notes**
Another definition states that process evaluation:  
…seeks to answer the question, "What services are actually being delivered and to whom?" Process evaluation includes such things as  
- tracking quantity and description of people who are reached by the program,  
- tracking quantity and types of services provided,  
- descriptions of how services are provided,  
- descriptions of what actually occurs while providing services, and  
- quality of services provided. (Van Marris and King 2007)

A more general definition of process evaluation is that it is the assessment of the m/w/s [merit/worth/significance] of everything that happens or applies before true outcomes emerge, especially the vision, design, planning, and operation of the program… (Scriven 2007)

The author of this review includes in process evaluations a review of key activity-supporting processes such as communication, relationship building, capacity building and resource management.

Sometimes the terms process and formative evaluations are used interchangeably.

**Outcome Evaluation**

**Handbook definition**
- An outcome evaluation examines the extent to which a project has achieved the outcomes it set at the outset.

**Notes**
An example of a slightly different definition follows:  
This type of evaluation determines whether the benefits or services to clients produced as outcomes of a program are meeting the needs of those clients… (Saskatchewan Industry and Resources 2004)

Outcome evaluations are sometimes called summative or impact evaluations but these terms cannot always be used interchangeably. According to some documents, impacts are the immediate effects of a program and outcomes are the more long-term results; others reverse the order. For example:

Impact is defined as the immediate effect that health promotion programs have on people, stakeholders and settings to influence the determinants of health…Outcome evaluation is linked to assessing the endpoint of interventions expressed as
### Summative Evaluation

**Handbook definition**
- Summative evaluations examine the overall effectiveness and impact of a project, its quality, and whether its ongoing cost can be sustained.

**Notes**
An example of another definition follows:

Summative evaluation focusses on programs that are already underway or completed. It investigates the effects of the program, both intended and unintended. It seeks to answer the questions “Did the program make a difference?” (impact evaluation) and “Did the program meet its stated goals and objectives?” (outcome evaluation). (Van Marris and King 2007)

### Economic

**Handbook definition**
- A cost-effectiveness study examines the relationship between project costs and project outcomes. It assesses the cost associated with each level of improvement in outcome.
- Cost-benefit analysis is like cost-effectiveness analysis in

**Notes**
Cost-effectiveness and cost-benefit evaluations are two different types of economic appraisal. A third type is cost utility analysis:

which evaluates the utility of different outcomes for different users or consumers of a policy or service. Cost utility analysis typically involves subjective appraisals and evaluations of outcomes using qualitative and quantitative data. (Government Social Research Unit 2007)
<table>
<thead>
<tr>
<th><strong>Handbook definition</strong></th>
<th><strong>Notes</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>that it looks at the relationship between project costs and outcomes (or benefits). But a cost-benefit study assigns a dollar value to the outcome or benefit so that a ratio can be obtained to show the number of dollars spent and the number of dollars saved.</td>
<td></td>
</tr>
</tbody>
</table>

**Meta-Evaluation**

<table>
<thead>
<tr>
<th><strong>Handbook definition</strong></th>
<th><strong>Notes</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>• [not defined]</strong></td>
<td>One definition of meta-evaluation follows: The term is used for evaluations designed to aggregate findings from a series of evaluations. It can also be used to denote the evaluation of an evaluation to judge its quality and/or assess the performance of the evaluators. (Development Assistance Committee [DAC] Working Party on Aid Evaluation 2002)</td>
</tr>
</tbody>
</table>

**Ongoing Review**

<table>
<thead>
<tr>
<th><strong>Handbook definition</strong></th>
<th><strong>Notes</strong></th>
</tr>
</thead>
</table>
| **• Project monitoring** counts specific project activities and operations. This is a very limited kind of evaluation that helps to monitor, but not assess the project. | Monitoring is sometimes called performance measurement: Performance measurement is the ongoing monitoring of the results of a program, policy or initiative, and in particular, progress towards pre-established goals. (Treasury Board of Canada 2001)

Some documents do not consider monitoring to be a type of evaluation:

- Evaluation should not be confused with monitoring. Monitoring is the systematic collection of routine financial and management information during the implementation of a project, programme or policy.
- It provides an essential source of information, indicating the extent to which objectives are being achieved. This regular review can give an early warning of potential problems.
- Monitoring also provides information which is useful at the evaluation stage. To be effective, plans for monitoring and evaluation should form part of the initial planning of a project. (Department of Health 2002) |
**Handbook definition**

In addition to information directly related to program or policy activities, processes and costs, some types of monitoring relate to information important to program or policy direction, such as demographics or social conditions (e.g. see Saskatchewan Department of Community Resources and Employment 2003)

Topics such as monitoring, performance management and quality assurance or improvement come up frequently in the evaluation literature.

One document defines performance management this way: …performance management may be construed generally as the process by which we collect performance information and use it for improvement, by drawing on a range of evaluative processes such as performance measurement and evaluation. (Davies 1999)

Some documents identify ongoing review activities such as performance management as a type of evaluation while others do not. Regardless of their position on whether ongoing review activities are part of evaluation, most documents agree on the importance of these activities to evaluation. For example, Davies (1999) notes that: …some jurisdictions see [evaluation and performance management] as having a mutually beneficial relationship, so that evaluation provides the framework to develop ‘good’ indicators for performance monitoring purposes, and performance data that are collected in an ongoing manner can be used for evaluation purposes.

Another document states:

In terms of evaluation, quality assurance is relevant in the sense that it supports the development of a systematic and reflective approach to improving practice and performance. It is also a critical component of process evaluation since in trying to understand and explain outcome patterns, it is important to be able to distinguish between the effects of the intervention itself and the quality of the delivery. (Wimbush and Watson 2000)

One document distinguishes between program review (similar to what many would consider typical evaluation), and Business Process Improvement (BPI) (along the lines of performance management or quality approaches):

While program review asks the fundamental questions: does this activity or program directly relate to our core strategic goals? Why do we do this? BPI delves within processes or programs and can dissect particular elements of a process or
program for improvement. BPI can be viewed as a complementary process to program review, albeit usually smaller in scope. (Saskatchewan Industry and Resources 2004)

Other Evaluation-Related Concepts

More definitions of evaluation-related concepts follow. Each definition is quoted verbatim from its source. Note that other documents may define these concepts in different ways. Concepts are listed alphabetically.

<table>
<thead>
<tr>
<th>Concept</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>baseline</td>
<td>Information collected before a project or activity begins is known as baseline information. It shows what the situation was like in the community or for individual participants before the project or activity began or before individual participants entered the project or activity. (Public Safety and Emergency Preparedness Canada 2006)</td>
</tr>
<tr>
<td>benchmarking</td>
<td>The literature distinguishes between standard benchmarking (assessment against a given performance standard), results benchmarking (comparing the performance of a large number of public organizations), and process benchmarking (comparing an organization's production process with that of an organization considered to be leading in the field — 'best practice'). (Hansen 2005)</td>
</tr>
</tbody>
</table>
| design    | Research designs fall into two main categories: experimental (including quasi-experimental) and observational (including analytic and descriptive). Sometimes designs are combined resulting in a less clear distinction... **Experimental designs** are often used to test efficacy or effectiveness of an intervention. (Efficacy is how well something works in an ideal situation and effectiveness is how well it works in a real-life situation.) Experimental designs involve:  
  – a planned intervention  
  – control of a relatively small number of selected factors (independent variables) which are analyzed with respect to their impact on or relationship to other factors (dependent variables)  
  – random assignment of study participants into intervention and control (comparison) groups which are compared after the intervention has been introduced  
  – (frequently) testing of a hypothesis (e.g. related to cause and effect)  
  The randomized control trial (RCT) is the major example of an experimental design... **Quasi-experimental designs** also involve a planned and controlled intervention but do not include randomization. They may or may not include a control group. The before and after study is a common type of quasi-experimental design... **Observational designs**, in contrast to experimental designs, do not involve a planned and controlled intervention. In an observational design, naturally occurring events or situations beyond the researchers' control are observed. |
<table>
<thead>
<tr>
<th><strong>Concept</strong></th>
<th><strong>Definition</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>data collection tools</strong></td>
<td>Methodologies used to identify information sources and collect information during an evaluation. Note: Examples are informal and formal surveys, direct and participatory observation, community interviews, focus groups, expert opinion, case studies, literature search. (Development Assistance Committee [DAC] Working Party on Aid Evaluation 2002)</td>
</tr>
</tbody>
</table>
| **ethical principles** | guidelines for appropriate values-based conduct (Kahan & Goodstadt 2005)  
*Note:* Methods to ensure informed consent and confidentiality of evaluation participants are a common ethical practice.) |
| **evaluability** | … the central requirements of evaluability…are that the interventions, and the target population, are clear and identifiable; that the outcomes are clear, specific and measurable; and that an appropriate evaluation design can be implemented (Patton, 1990). (Government Social Research Unit 2007) |
| **goals** | what we strive to achieve; the concrete manifestation of values (Kahan & Goodstadt 2005) |
| **indicator** | A variable (or information) that measures one aspect of a program or project. It indicates whether a project has met a particular goal. There should be at least one indicator for each significant element of the project (i.e., at least one for each outcome identified in the logic model). …There are two kinds of indicators:  
– A process indicator provides evidence that a project activity has taken place as planned.  
– An outcome indicator provides evidence that a project activity has caused a change or difference in a behaviour, attitude, community, etc. (Public Safety and Emergency Preparedness Canada 2006) |
| **informed consent** | An ethical requirement where participants give permission for the sharing of their information and experiences. This usually involves signed agreements which are intended to protect the participants and guarantee their anonymity. (Farell et al. 2002) |
| **input** | The financial, human, and material resources used for the development intervention. (Development Assistance Committee [DAC] Working Party on Aid Evaluation 2002) |
| **logic model or logical framework** | A logic model is a way of describing a project or program. It is a tool to help in project planning and evaluation. A logic model describes the resources and activities that contribute to a project and the logical links that lead from project activities to the project’s expected outcomes. Logic models are often depicted as a flow chart that includes the project’s inputs, activities, outputs, and outcomes. (Public Safety and Emergency Preparedness Canada 2006) |
| **methods** | Methods involve data collection and data analysis, which can be either quantitative or qualitative in nature. Some studies use both quantitative and qualitative methods. The data that are collected, and the ways in which they are analyzed, can also be quantitative or qualitative in nature. (Kahan &
<table>
<thead>
<tr>
<th>Concept</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>objectives</strong></td>
<td>• what we need to accomplish in order to achieve a more general goal; objectives are generally designed to be observable or measurable (Kahan &amp; Goodstadt 2005)</td>
</tr>
<tr>
<td></td>
<td>(Note: Objectives can be outcome objectives, process objectives, or activity objectives.)</td>
</tr>
<tr>
<td><strong>outputs</strong></td>
<td>• Outputs are the services or products you will develop and provide. (Van Marris &amp; King 2007)</td>
</tr>
<tr>
<td></td>
<td>• Process evaluation may also include the evaluation of what are often called “outputs,” (usually taken to be “intermediate outcomes” that are developed en route to “true outcomes,” a.k.a. longer term results or impact) such as knowledge, skill, or attitude changes in staff (or clients), when these changes are not major outcomes in their own right. (Scriven 2007)</td>
</tr>
<tr>
<td></td>
<td>• The products, capital goods and services which result from a development intervention; may also include changes resulting from the intervention which are relevant to the achievement of outcomes. (Development Assistance Committee [DAC] Working Party on Aid Evaluation 2002)</td>
</tr>
<tr>
<td></td>
<td>(Note: Not all evaluations use the term “outputs.” What one evaluation calls an output, another might call: an activity; an indicator of whether a process or activity objective has been achieved; or a short-term outcome.)</td>
</tr>
<tr>
<td><strong>performance</strong></td>
<td>• Performance means actual achievements measured against defined goals, standards or criteria. (Treasury Board of Canada 2001)</td>
</tr>
<tr>
<td><strong>policy</strong></td>
<td>• A principle, plan, or course of action that is put into place or carried out by governments, organizations, or individuals. (Farrell et al. 2002)</td>
</tr>
<tr>
<td><strong>project or program objective</strong></td>
<td>• The intended physical, financial, institutional, social, environmental, or other development results to which a project or program is expected to contribute. (Development Assistance Committee [DAC] Working Party on Aid Evaluation 2002)</td>
</tr>
<tr>
<td><strong>program</strong></td>
<td>• A program is any group of related, complementary activities intended to achieve specific outcomes or results. (Van Marris &amp; King 2007)</td>
</tr>
<tr>
<td><strong>qualitative</strong></td>
<td>• [Examples of] Qualitative methods:</td>
</tr>
<tr>
<td></td>
<td>– Focus groups</td>
</tr>
<tr>
<td></td>
<td>– In-depth interviews</td>
</tr>
<tr>
<td></td>
<td>– Open-ended survey questions</td>
</tr>
<tr>
<td></td>
<td>– Diaries</td>
</tr>
<tr>
<td></td>
<td>– Consensus building (Delphi Method)</td>
</tr>
<tr>
<td></td>
<td>– Forums/discussion groups (Van Marris &amp; King 2007)</td>
</tr>
<tr>
<td></td>
<td>• Qualitative data include:</td>
</tr>
<tr>
<td></td>
<td>– nominal data: any items that can be categorized (e.g. a person with characteristics a, b and c is “male” and a person with characteristics d, e and f is “female”)</td>
</tr>
<tr>
<td></td>
<td>– ordinal data: items that are ranked (e.g. the worst and best pens) (Kahan &amp; Goodstadt 2005)</td>
</tr>
</tbody>
</table>
### Concept  |  Definition
--- | ---
**quantitative** | • [Examples of] Quantitative methods:  
  – Intercept, mail or telephone survey  
  – Process tracking forms/records  
  – Service utilization  
  – Analysis of large datasets  
  – Direct measures of health indicators/behaviours (e.g., blood pressure)  
  – Direct measures of illness (morbidity or mortality rates) (Van Marris & King 2007)  
  • Quantitative data include:  
    – interval data: have no true zero and so there can be no ratio between two numbers (e.g. 20 degrees Celsius is not twice as warm as 10 degrees Celsius)  
    – ratio data: can be expressed as a ratio (e.g. the income for person or group “a” is only half the income of person or group “b”, that is, a ratio of one to two) (Kahan & Goodstadt 2005)

**reliability** | • The extent to which any measuring device yields the same results each time it is applied to a population or program (Farell et al. 2002)

**results** | • The output, outcome or impact (intended or unintended, positive and/or negative) of a development intervention. (Development Assistance Committee [DAC] Working Party on Aid Evaluation 2002)

**stakeholder** | • Who are the stakeholders? That varies from project to project, depending on the focus, the funding, the intended outcomes, etc. There are a number of groups that are generally involved, however:  
  1. Participants or beneficiaries  
  2. Project line staff and/or volunteers  
  3. Administrators  
  4. Outside evaluators, if they're involved.  
  5. Community officials.  
  6. Others whose lives are affected by the project (Community Tool Box undated)

**validity** | • The extent to which a test actually measures what it is intended to measure.* (Farell et al. 2002)

**values** | • what is most important to us (Kahan & Goodstadt 2005)
Overview of Approaches

Many evaluation approaches exist. The literature suggests that no one approach is best for all situations. Rather, the best approach varies according to factors such as fit with basic values, the intent of the evaluation, the nature of key stakeholders, and available resources. For example:

The approach you choose for your evaluation will depend on the evaluation’s purpose. If you wish to learn ways to improve the services you offer, a utilization-focused or an empowerment approach might be appropriate. If you want to convince outside organizations that you are having a positive impact on participants, an external evaluator will help to assure objectivity. (Public Safety and Emergency Preparedness Canada 2006)

In addition, it is not necessary to stick strictly to one approach: evaluations “might quite reasonably combine elements of different approaches or adapt to local conditions.” (Rogers & Fraser 2003)

Note that most evaluation approaches can be applied to any type of evaluation. For example, a theory based, collaborative or utilization-focused approach can be used to conduct either a process or outcome evaluation.

Steps to Guide Evaluation Practice

Each approach has a number of associated sets of steps which guide the practice of evaluation. Regardless of the approach there is a large degree of overlap in the suggested steps. Steps related to any particular approach will differ in the nature of the methods and tasks related to each step. Many descriptions of steps emphasize their iterative nature and that a particular order is not always followed.

An example of a fairly typical set of steps follows, from Van Marris and King (2007). It differs from other sets of steps in two ways. First, it is more comprehensive than many and is therefore useful as a checklist to ensure no major steps are omitted. Second, the step “engage stakeholders” is relevant to collaborative evaluation approaches but may not be relevant to other approaches. Although designed to evaluate health promotion programs, the set of steps below is generalizable to many other types of programs and policies.

1. Clarify your Program: e.g. define goals, population of interest, outcome objectives, activities, measurable indicators
2. Engage Stakeholders
3. Assess Resources for the Evaluation: e.g. staff, funding
4. Design the Evaluation: e.g. select evaluation type and framework, consider ethical issues and confidentiality
5. **Determine Appropriate Methods of Measurement and Procedures**
6. **Develop Work Plan, Budget and Timeline for Evaluation**
7. **Collect the Data Using Agreed-upon Methods and Procedures**
8. **Process and Analyze the Data**
9. **Interpret and Disseminate the Results**
10. **Take Action**

Figure 1 below provides a comprehensive set of steps from Darabi (2002). It provides a pictorial overview of the relationships between the steps. Note the inclusion of standards of utility, feasibility, propriety and accuracy mentioned in the section *Guiding Principles*. *(Note: AEA in the figure stands for American Evaluation Association.)*
Figure 1. The Phases and Components of a Systems Approach to Conducting Program Evaluation.
Summary of Major Evaluation Approaches

A summary of major approaches, based on the literature, follows. Approaches vary on the basis of what is evaluated, who participates in the evaluation, evaluation purpose, and how the evaluation is conducted. In practice, approaches are often combined.

Table 2: Summary of Major Evaluation Approaches

<table>
<thead>
<tr>
<th>Description</th>
<th>Strengths</th>
<th>Cautions and other notes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Goals Based</strong></td>
<td>The evaluation identifies whether program/policy goals/objectives are met.</td>
<td>• Determines whether results align with goals/objectives, important for accountability reasons and to identify change.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Goal Free</strong></td>
<td>The evaluation identifies all results of the program, whether anticipated or unanticipated.</td>
<td>• More likely to identify unanticipated results.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Theory Based</strong>&lt;br&gt;(logic model)</td>
<td>The evaluation is guided by the program/policy theory, i.e. the logical relationships between all the parts.</td>
<td>• Encourages a greater understanding of the fundamental mechanisms of the program/policy, i.e. what works when, where, why and for whom.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Decreases the chance of negative results.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Increases dialogue among a variety of stakeholder groups and clarifies underlying assumptions.</td>
</tr>
<tr>
<td><strong>Utilization</strong></td>
<td>Evaluation is structured according to who will use the evaluation results and how they will use the results.</td>
<td>• Pragmatic and flexible.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Increases chances evaluation results will be followed up on.</td>
</tr>
<tr>
<td><strong>Collaborative</strong></td>
<td>Collaborative evaluation is a broad</td>
<td>• Strong capacity building potential.</td>
</tr>
<tr>
<td>Description</td>
<td>Strengths</td>
<td>Cautions and other notes</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------</td>
<td>--------------------------</td>
</tr>
</tbody>
</table>
| category which includes a variety of types of evaluation, with varying degrees of participation and different purposes. | • Increases relevance of results.  
• Increases chances of results being followed up on because active participants feel more invested in the evaluation.  
• Greater range of perspectives represented in evaluation information.  
• May give voice to traditionally marginalized groups. | • Time consuming.  
• May require larger amounts of funds.  
• May lack credibility to some.  
• Requires strong facilitation because of diverse perspectives and situations.  
• Attaining equitable participation is challenging because people’s circumstances and degree of power vary so much. |

**Balanced Scorecard**

The Balanced Scorecard approach to evaluation is based on four linked areas: financial, customer, business process, learning and growth.

| | Systematic.  
• Increases clarity and consensus around strategy.  
• Increases alignment between strategic objectives and actions. | Top-down approach.  
• Possibility of faulty assumptions leading to negative results if a causal link between the perspectives is assumed. |

**Appreciative Inquiry (AI)**

Appreciative Inquiry uses an asset-based approach to evaluation.

| | Increases evaluation participation.  
• Increases use of results.  
• Builds learning and change capacity.  
• Supports and strengthens good performance. | A debate exists regarding whether AI avoids addressing challenges.  
It is not always appropriate.  
Requires a special skill set to carry out well.  
Program staff are more likely to be included than clients, increasing the power gap. |

**External**

An evaluator not associated with the organization or program carries out the evaluation.

| | Greater credibility.  
• Efficient. | May be irrelevant to stakeholders.  
May miss or misinterpret information. |

**Kirkpatrick**

Kirkpatrick’s approach is to evaluate four levels sequentially: reaction, learning, behaviour, results.

| | Straightforward. | Does not explore the why or how of results. |

**CIPP**

CIPP evaluations focus on program context, input, process, and product.

| | Takes into account a range of environmental factors from politics to personalities. | Time consuming. |

More detail about each major approach follows. Examples of application steps are provided for each approach. Note that step descriptions are brief; some of the original documents from which the steps are taken run to 100 pages or more. For this reason it is important to view the entire document for a complete understanding of the steps.
Results-Focused Evaluation

Results-focused evaluations look at the difference a policy or program has made to individuals, groups or the broader community. The types of results-focused evaluations discussed below are: goals based; goal free; combined goals based and goal free; and theory based.

**Goals Based**

In the goal-attainment model, which is the classic model in the literature on programme evaluation and organizational evaluation, results are assessed only in relation to the predetermined goals. The evaluator closely scrutinizes the object of evaluation in relation to its goals. (Hansen 2005)

The term “goals” in this approach is used broadly to include objectives, performance targets and expected outcomes. The goals-based approach to evaluation was developed by Tyler in the late 1940s and has continued to evolve since. For example, an adaptation in the 1970s was:

- to determine why goals are attained including a three-dimensional model: instruction (characteristics of the program or activity), institution (characteristics of individuals or groups), and behavior (cognitive, affective, psychomotor). (McEwen 2006, citing Hammond 1973)

The key strength of this model is the determination of whether results align with goals. The degree of alignment between goals and results is important to know from accountability and improvement perspectives.

Hansen (2005) lists key requirements for the successful use of this approach:

- Cameron (1986) argues that using the goal-attainment model is only possible and useful if the goals of the evaluated object are clear, if there is consensus about them, and if they are time-bound and measurable.

Cautions regarding use of this approach follow.

- This approach “disregards side effects and unanticipated effects; it does not consider costs; and the relevance of the goals is not questioned.” (Mickwitz 2003)
- The approach does not question whether “the chosen goals or targets are valid or appropriate measures of effectiveness.” (Government Social Research Unit 2007)
- If this approach is seen as equivalent to an outcome or summative evaluation, underlying mechanisms and process questions might be ignored.
- A match between outcomes and goals is not necessarily the result of the policy or program. “Other factors, including other policy initiatives, may have been responsible.” (Government Social Research Unit 2007)
Goals Free

Policy makers and evaluators are often interested in the unintended consequences or outcomes of a policy, programme or project. These unintended outcomes may be beneficial or harmful. Goals-free evaluation does this by focusing on the actual effects or outcomes of some policy, programme or project, without necessarily knowing what the intended goals might be. (Government Social Research Unit 2007)

In the original form of this approach, developed by Scriven in the 1970s, the evaluator was not told the goals or objectives of the program in order to identify all important unanticipated results. According to Walden & Baxter (2001), “Although this model was never put into practice, the concept of spin-offs [unplanned side effects] has lived on.”

The key strength of this type of evaluation approach is that it looks for unanticipated results. In support of goals-free evaluation, one document argues that goals-free evaluation:

should be of interest to government social researchers and policy analysts because of the importance of establishing the balance between the positive and negative consequences of policies. Such a balanced evaluation is important in order to establish the cost-benefit and cost-utility of a policy or programme intervention.” (Government Social Research Unit 2007)

Cautions regarding use of this approach follow:

- It may have “deficient evaluative criteria. The evaluator may circumvent this problem by applying the model to a before/after analysis, i.e. to analyse the organizational field before and after the given performance, programme or organization was implemented.” (Hansen 2005)
  
  The reference to deficient evaluative criteria refers to the view that “Having explicit standards for judgment is a defining attribute of evaluation that distinguishes it from other approaches to strategic management.” (Milstein et al. 2000)

- Similar to the goals-based approach, this approach “leaves the cost criticism unanswered.” (Mickwitz 2003)

- Principles of accountability may require an examination of whether policies or programs meet their goals, which this approach does not. “Another critique against the goal-free evaluation model is that if goals set by elected political bodies are not followed up, this can hardly be in line with a transparent democratic process (Vedung, 1997: 61).” (Mickwitz 2003)
Goals Based and Goals Free Combined

It is fairly common to combine the two approaches. That is, evaluations examine whether goals are met; they also look for unanticipated results, whether positive or negative. For example:

One evaluation model that is often particularly suitable for the evaluation of environmental policy instruments is ‘side-effects evaluation’ (Vedung, 1997). This is due to both the complexity and the uncertainty of many environmental policy problems. In this model the effects of the instruments evaluated are first divided into anticipated and unanticipated effects. The next level examines whether the effects occur inside or outside the target area. The third level is a qualitative categorization of the effects. (Mickwitz 2003)

Steps: Example 1

An example of steps to develop and implement a goals-based evaluation approach which includes the collection of unexpected data follows. This information is from Interagency Projects: An Evaluation Guide (ADM’s Forum on Human Services 1997). According to this guide, “The primary reason for evaluating programs is to determine the extent to which the program objectives have been achieved.” It focuses on expected outcomes but also encourages examination of unexpected data. The guide discusses implementation evaluation in addition to outcomes evaluation. Most sets of evaluation steps mix approaches to some extent.

1. Getting Ready: e.g. set up an evaluation committee; describe the program, including identifying program objectives and expected outcomes
2. Planning: e.g. identify evaluation purpose, users, questions, indicators, data sources, data collection methods, evaluation team members roles and responsibilities, ethical guidelines; develop a schedule
3. Implement the Evaluation Plan: e.g. monitor responses to information requests, schedule, and evaluation team roles and responsibilities; change the evaluation plan as necessary
4. Analyze Data and Draw Conclusions: e.g. make conclusions based on the data
5. Report the Conclusions: e.g. consider the audience, use tables and diagrams, include a summary

The general steps outlined above are common to many approaches. As mentioned before, differences are often in the details and emphasis. For example, not all approaches would place as strong an emphasis on the development of program objectives and indicators.

Theory Based Evaluation

The programme theory model may be seen as an extended results model. Where the classical results models focus only on concrete goals or effects related to a specific intervention or organization, and aims to evaluate the results of these,
the programme theory model opens up the underlying black box of the programme theory, uncovers mechanisms and raises the focus to a cluster of interventions or to an organizational field. The aim of the programme theory model is to revise and further develop programme theory and thus learn what works for whom in which contexts. (Hansen 2005)

One key informant talked about the social science, educational, health promotion or other theory driving a policy or program. For example, if the underlying theory relates to learning: “You might have some steps where people read literature and from that change their attitudes and as a result of changed attitudes they change behaviour.”

Sometimes theories which drive a program or policy are not explicit. A theory-based evaluation will identify underlying theories by reviewing program details and then examine whether this theory works in practice. Evaluation results will provide insights about required changes either to the theory or to the processes and activities designed to achieve the desired outcomes. In other words, the evaluation will provide a greater understanding of how to reach particular policy or program goals.

Often a logic model is used to summarize in visual form the underlying theory of the policy or program by highlighting the internal linkages among the various parts of the program, for example processes, activities and outcomes. For this reason theory-based evaluation approaches are sometimes called Program Logic or Logical Framework approaches. In the same way, “program theory” and “logic model” often refer to the same thing.

Evalulators often start out by clarifying a programme’s aims, objectives and the desired outcomes, but theory-based approaches suggest that evaluators also go on to elicit the key assumptions and linkages underlying how a programme has been designed, i.e. understanding the ‘logic’ of how the programme is supposed to operate to achieve the desired outcomes. (Wimbush & Watson 2000)

Below is an example of a logical model from Sullivan et al. (2007):
A strong rationale for using theory-based evaluation is that:

Failure to be clear about the causal sequence by which a policy is expected to work can result in well intentioned policies being misplaced, and outcomes that are contrary to those that were anticipated. Theory-Based evaluation provides a number of ways of carrying out an analysis of the logical or theoretical consequences of a policy, and can increase the likelihood of the desired outcome being achieved. (Government Social Research Unit 2007)

Other benefits result from the development and use of logic models:

Effective evaluation and program success rely on the fundamentals of clear stakeholder assumptions and expectations about how and why a program will solve a particular problem, generate new possibilities, and make the most of valuable assets. The logic model approach helps create shared understanding of and focus on program goals and methodology, relating activities to projected outcomes. (W.K. Kellogg Foundation 2004)
Although many documents write favourably of a theory-based approach, not everyone totally supports this approach:

...if you can be helpful by evaluating the program theory, do it; but it’s not an essential part of doing a good evaluation and will often be a diversion and sometimes a cause for antagonism. (Scriven 2007)

Other potential limitations are the time required to identify the theory or develop the logic model and the possibility that a logic model “does not always capture all aspects of the program (e.g. program costs may not be included in the model).” (Farell et al. 2002)

Another concern mentioned in the literature is that:

The matrix [of a Logical Framework Approach] lays heavy emphasis on objective achievement and goal attainment and the emphasis is often on measurable outcomes; there is less room for qualitative outputs... (Walden & Baxter 2001)

**Steps: Example 2**
The set of steps below is from Saunders et al. (2005). The steps focus on planning a process evaluation, which “looks inside the so-called black box to see what happened in the program and how that could affect program impacts or outcomes.”

1. **Describe the Program:** e.g. purpose, underlying theory, objectives, strategies, expected outcomes; “Ideally, this should be conveyed in a logic model that specifies the theoretical constructs of interest, those expected to change, and mediators of the change process (Scheirer et al., 1995; Steckler & Linnan, 2002a).”

2. **Describe Complete & Acceptable Program Delivery:** “A description of complete and acceptable delivery of the program should be based on the details of the program (e.g., program components, theory, and elements in logic model) and guided by an external framework such as the recommended elements of a process evaluation plan” which include fidelity, dose, and reach. Fidelity refers to the “quality of implementation”; dose refers to the “amount of program delivered by implementers” and the “extent to which participants receive and use materials or other resources” (i.e. frequency and duration); and reach refers to the “degree to which the intended priority audience participates in the intervention.” (Note that the term “intervention” is often interchangeable with other terms such as program, activity, policy or initiative.)

3. **Develop Potential List Of Questions:** This step includes identification of “the information needed to answer each process-evaluation question.”

4. **Determine Methods:** “Primary issues to consider in planning process-evaluation methods include design (when data are to be collected), data
sources (from where the information will come), tools or measures needed to collect data, data-collection procedures, data-management strategies, and data-analysis or data-synthesis plans...”

5. **Consider Program Resources, Context & Characteristics:** e.g. organizational systems and structures, external environment factors such as political factors, program’s age, size and complexity

6. **Finalize the Process Evaluation Plan**

The figure below (Saunders et al. 2005) illustrates these steps:

*Figure 3: Process Evaluation Steps (Saunders et al. 2005)*

This set of steps strongly emphasizes program theory:

Having a well-planned and theory-based health-promotion program is the beginning point for planning process evaluation. The approach to process-evaluation planning described in this article assumes that the intervention has been planned in detail with guidance from appropriate theories and/or a conceptual model. (Saunders et al. 2005)

Note that because the steps were designed for the development of an evaluation plan, not for the implementation of the evaluation itself, steps such as collecting and analyzing data and reporting the results are omitted.
Steps: Example 3

The following example is of the Health Education Board for Scotland (HEBS) Evaluation Framework, from Wimbush & Watson (2000):

The HEBS framework identifies the different purposes of evaluation and the associated evaluation questions that are characteristic of each of these stages, acknowledging the importance of assessing effectiveness, as well as assuring quality and making explicit the mechanisms of change implicit in a programme’s theory.

In this example, provided in Figure 4, the framework outlines the stages in the development of a program and type of evaluation most appropriate for each stage, along with evaluation purpose, research questions, application, and responsibility and resources.
### Figure 4: HEBS Evaluation Framework (Wimbush & Watson 2000)

<table>
<thead>
<tr>
<th>Stage of project development</th>
<th>Type of evaluation</th>
<th>Purpose of evaluation</th>
<th>Research questions</th>
<th>Application</th>
<th>Whose responsibility and what resources are required?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planning</td>
<td>Systematic reviews of effectiveness.</td>
<td>Overview of evidence of effectiveness from outcome evaluations around a specific topic setting.</td>
<td>What are effective ways of addressing a particular need or problem? What makes a difference?</td>
<td>All (where possible).</td>
<td>International collaborations, government departments and national agencies.</td>
</tr>
<tr>
<td>Design and pilot</td>
<td>Developmental evaluation.</td>
<td>To assess the feasibility, practicability and acceptability of the new project and its processes /mechanisms and to test the potential effectiveness of a new approach.</td>
<td>Is it feasible and practicable?</td>
<td>Projects in design or pilot stage which are testing new or innovative approaches.</td>
<td>Funders/sponsors of local programmes.</td>
</tr>
<tr>
<td>Implementation – early start up</td>
<td>Monitoring and review (for evaluation and quality assurance).</td>
<td>To monitor and review progress in achieving agreed milestones and agreed quality standards in order to improve quality and efficiency.</td>
<td>What have we achieved so far? How could we improve?</td>
<td>All.</td>
<td>Project managers, Cost of training in project management and quality assurance which should include review and monitoring procedures.</td>
</tr>
<tr>
<td>Implementation – establishment</td>
<td>Impact evaluation.</td>
<td>To assess the short-term effectiveness of a project in terms of its reach and immediate impacts. If data on costs are available, simple economic evaluation measures can be produced.</td>
<td>How is the project working? Is it being implemented as intended? To what extent is the target population being reached? To what extent are objectives/goals being achieved? What impacts have been achieved? At what cost?</td>
<td>All large projects (total budget £100k+).</td>
<td>Funders/sponsors of local programmes. Funding for evaluation should be proportional to project costs. Guidelines suggest approx. 10–15% of total project costs.</td>
</tr>
<tr>
<td>Implementation – fully operational</td>
<td>Outcome evaluation.</td>
<td>To assess the longer-term effectiveness of a project using intermediate outcome measures. If data on costs are available, more complex economic evaluation measures can be produced.</td>
<td>To what extent were intermediate outcomes achieved? How were these achieved? In which groups/settings are the greatest benefits shown?</td>
<td>Projects where short-term effectiveness has been assessed positively.</td>
<td>Source of funding is often research council grants, trust funds, or government departments. Applications are led by academic researchers with evaluation expertise, but developed in collaboration with programme team.</td>
</tr>
<tr>
<td>Dissemination</td>
<td>Transfer evaluation.</td>
<td>To assess the replicability of a project's mechanisms/processes and outcomes.</td>
<td>Can the project be transferred to another setting or population and achieve the same outcomes?</td>
<td>All large projects that have proved effective in one setting/population.</td>
<td>Local agencies who want to apply approach locally. Government agencies wanting to disseminate effective practice.</td>
</tr>
</tbody>
</table>
Steps: Example 4

The Interactive Domain Model (IDM) best practices framework was designed as a planning and evaluation tool to operationalize the IDM approach for better health. The IDM is based on the theory that best practices are context-sensitive and occur, when there is consistency amongst the following domains (Kahan & Goodstadt 2001):

- underpinnings: values, goals and ethics; theories and beliefs; evidence
- understanding of the environment (physical, psychological, social, political, economic)
- practice: processes and activities related to addressing selected health/social or organizational issues and conducting research and evaluation

The IDM proposes four key planning and evaluation steps, each of which is applied to underpinnings, understanding of the environment, and practice (Kahan & Goodstadt 2005):

1. **Prepare Foundation for Action re. selected issue**: What guides us? Where are we now? Where do we want to go?
2. **Make Action & Evaluation Plan**: How do we get to where we want to go? Who does what, when & how? This step includes identifying specific objectives to achieve the ideal, resources, and challenges, and developing an evaluation plan.
3. **Implement, Reflect & Document**: What did we do? How did we do it? What were the results?
4. **Revise**: What do we need to change?

The IDM Framework is shown below.
The IDM Evidence Framework provides more detailed steps for guiding an evaluation. A summary of these steps follows (Kahan & Goodstadt 2005).

- **Set the Stage**: e.g. identify: general guidelines; the issue; priority questions; key stakeholders and their roles; processes to deal with differences of opinion/perspective; cultural/other sensitivities; system for tracking findings/decisions

- **Generate Information**: e.g. identify design and methods; identify and select information sources according to a set of criteria such as quality and applicability to specific context

- **Analyze Data**: e.g. identify: methods used; outcomes; relationship between outcomes, activities and processes; conclusions. This step includes summarizing or synthesizing information: e.g. assess transferability of findings; identify relevant learnings and limitations.

- **Report On Findings**: e.g. identify audience, ensure high quality reporting

- **Make Decisions Based On Results**: implement decisions; evaluate; revise

### Figure 5: IDM Framework (Kahan & Goodstadt 2005)

The Interactive Domain Model (IDM) Framework for Best Practices in Health Promotion & Public Health
Barbara Kahan and Michael Goodstadt (May 2005, revised from 2001 version)
Utilization-Focused Evaluation

Utilization-Focused Evaluation (U-FE) begins with the premise that evaluations should be judged by their utility and actual use; therefore, evaluators should facilitate the evaluation process and design any evaluation with careful consideration of how everything that is done, from beginning to end, will affect use. Use concerns how real people in the real world apply evaluation findings and experience the evaluation process. Therefore, the focus in utilization-focused evaluation is on intended use by intended users. Since no evaluation can be value-free, utilization-focused evaluation answers the question of whose values will frame the evaluation by working with clearly identified, primary intended users who have responsibility to apply evaluation findings and implement recommendations. (Patton 2002)

Utilization-Focused Evaluation is associated with Michael Quinn Patton, quoted above. This approach is open ended:

Utilization-focused evaluation does not advocate any particular evaluation content, model, method, theory, or even use. Rather, it is a process for helping primary intended users select the most appropriate content, model, methods, theory, and uses for their particular situation. Situational responsiveness guides the interactive process between evaluator and primary intended users. A utilization-focused evaluation can include any evaluative purpose (formative, summative, developmental), any kind of data (quantitative, qualitative, mixed), any kind of design (e.g., naturalistic, experimental), and any kind of focus (processes, outcomes, impacts, costs, and cost-benefit, among many possibilities). Utilization-focused evaluation is a process for making decisions about these issues in collaboration with an identified group of primary users focusing on their intended uses of evaluation. (Patton 2002)

The obvious strengths of this approach are its pragmatic flexibility and the greatly increased likelihood that evaluation results will be followed up on.

One concern mentioned in the literature is that by focusing the evaluation on the interests of the intended users of the results, the role of other groups may be decreased. (Mickwitz 2003)

Steps: Example 5

Patton (2002) presents a list of “Primary U-FE tasks.” He comments that in working with the intended users “the evaluator offers a menu of possibilities within the framework of established evaluation standards and principles.” This list follows.

1. Program/Organizational Readiness Assessment
2. Evaluator Readiness and Capability Assessment
3. **Identification of Primary Intended Users**
4. **Situational Analysis**
5. **Identification of Primary Intended Uses:** e.g.: Consider how evaluation could contribute to program improvement, making major decisions, generating knowledge
6. **Focusing the Evaluation**
7. **Evaluation Design:** e.g.: select methods appropriate to the questions; “assure that results obtained by the methods selected will be believable, credible, and valid to primary intended users”; “assure that the proposed methods and measurements are practical, cost-effective, ethical”
8. **Simulation of Use:** “Before data are collected, a simulation of potential use can be done with fabricated findings in a real-enough way to provide a meaningful learning experience for primary intended users.”
9. **Data Collection**
10. **Data Analysis**
11. **Facilitation of Use**
12. **Metaevaluation**

**Collaborative Evaluation**

Three distinguishable features are used to characterize collaborative evaluation: control of the evaluation process (ranging from control of decision being completely in the hands of the evaluator or researcher to control being exerted entirely by practitioners); stakeholder selection for participation (ranging from restriction to primary users to inclusion of all legitimate groups); and depth of participation (from consultation—with no decision making control or responsibility—to deep participation— involvement in all aspects of an evaluation from design, data collection, analysis, and reporting to decisions about dissemination of results and use; Cousins & Whitmore, 1998). The evaluation is located somewhere on each continuum depending on who controls the process, who participates, and how much. (Butterfoss et al. 2001)

Many types of evaluation are classified as “collaborative,” ranging from democratic evaluation to empowerment evaluation. Butterfoss et al. (2001) provide an overview of the various types of collaborative evaluation:

- **practical participatory evaluation (P-PE):** “fosters evaluation use and assumes that evaluation is geared toward program, policy or organizational decision making. The central premise of P-PE is that stakeholder participation in evaluation will enhance relevance, ownership, and utilization.”
- **transformative participatory evaluation (T-PE):** “aims to empower people through participation in the process of respecting their knowledge (based on Friere’s, 1970, conscientization) and understanding the connection among knowledge, power, and control. Evaluators and participants work collectively, and participants are urged to consider their own biases and assumptions through critical reflection. T-PE has as its
primary function the empowerment of individuals and groups...In this process, evaluation methods and products are used to transform power relations and to promote social action and change.” Facilitators and participants make joint decisions about the evaluation.

- **stakeholder-based evaluation:** “similar to P-PE but is used in situations where stakeholder groups do not generally agree on program goals and need to be guided to consensus by the evaluator (Cousins & Earl, 1992). Thus, stakeholder-based evaluation is used best in decision-oriented, problem-solving, or summative evaluations (Mark & Shotland, 1985). Stakeholder-based evaluation differs from T-PE because of its practical goals, evaluator control, and limited stakeholder participation.”

- **democratic evaluation:** “similar to stakeholder-based evaluation because it increases the use of evaluation by having evaluators and participants share the work and decisions (MacDonald, 1976). It is democratic because participants have control over the interpretation and publication of data and results (McTaggart, 1991). Finally, representation from diverse stakeholder groups is an essential element of democratic evaluation.”

- **participatory action research:** “stakeholders who are members of the target organization participate both as subjects of research and as co-researchers (Argyris & Schon, 1991; Whyte, 1991). The goal of PAR is to inform and improve practice within organizations and to help them to change. Because researchers and organization members are partners, organization members are encouraged to be active in all research and evaluation activities.”

- **empowerment evaluation:** “uses evaluation concepts and techniques to foster self-determination. This type of evaluation is flexible, collaborative, and focuses on program improvement. The key objectives are training, facilitation, empowerment of individuals and groups, advocacy, and illumination of issues of concern (Fetterman, 1994; Fetterman, Kafterian, & Wandersman, 1996). Although this approach is closely linked to T-PE, critics have stated that empowerment evaluation does not exhibit enough emancipatory results and is in fact closer to P-PE in nature (Patton, 1997).”

While Butterfoss et al. describe six types of collaborative evaluation, often only two distinctions are made, between participatory and empowerment evaluation. Public Safety and Emergency Preparedness Canada (2006) provides a description for each:

[Participatory evaluation] is a method that involves participants in all aspects of the evaluation, from identifying the evaluation questions to deciding what information to collect, how to do it, and how to interpret the findings.

[Empowerment evaluation] is an approach that uses evaluation concepts, techniques, and findings to help community groups improve their programs and services. The evaluator acts as a coach or facilitator to help project staff and participants through a process of self-evaluation and reflection. Empowerment evaluation follows three steps: a) establishing a mission or vision statement, b)
identifying and prioritizing the most significant program activities and rating how well the program is doing in each of those activities, and c) planning strategies to achieve future project improvement goals (Fetterman, 2002).

Not all documents define the different types of collaborative evaluation the same way that Butterfoss et al. do. For example, the following definition refers to a “deliberative democratic perspective” which appears to combine Butterfoss et al.’s definitions of “stakeholder-based” and “democratic” evaluations:

[Deliberative democratic] evaluation incorporates democratic processes within the evaluation to secure better conclusions. The aspiration is to construct valid conclusions where there are conflicting views. The approach extends impartiality by including relevant interests, values, and views so that conclusions can be unbiased in value as well as factual aspects. Relevant value positions are included, but are subject to criticism the way other findings are. Not all value claims are equally defensible. The evaluator is still responsible for unbiased data collection, analysis, and arriving at sound conclusions. The guiding principles are inclusion, dialogue, and deliberation, which work in tandem with the professional canons of research validity. (House & Howe 2000)

The literature identifies many strengths related to one or more types of collaborative evaluation. Strengths most frequently mentioned are:

- capacity building
- greater breadth and depth of information because of the range of perspectives involved
- providing a voice to people who are often voiceless
- greater relevance to stakeholders with increased likelihood of follow up to results
- improvements to programs or social conditions

Butterfoss et al. (2001) add one qualification:

Evidence indicates that stakeholder participation will improve the quality of evaluation results, but the particular circumstances and conditions must be carefully considered.

Butterfoss et al. (2001) also cite a number of other cautions regarding collaborative approaches to evaluation:

Some disadvantages are that (a) extra time is needed to train and build relationships that are empowering; (b) participants’ motivation, commitment, and skills vary; (c) highly technical reports are not usually produced; (d) front-end costs are high; and (e) some rigor may be lost. Rigor may be improved as participants’ skills and commitment increase and by the use of systematic approaches that rely on multiple methods. (Butterfoss et al. 2001)
The amount of time required is probably the most common concern mentioned in the literature. Other concerns are:

- the need for “good participatory processes in place to facilitate diverse groups working together (e.g. decision making, communication, conflict resolution)” (Kahan & Goodstadt 2005)
- the requirement for “high levels of interpersonal and organizational trust” (Butterfoss et al. 2001)
- the importance of a “safe atmosphere” to allow empowerment evaluation participants to “enjoy the latitude to take risks and simultaneously assume responsibility” (Townsend & Adams 2003)
- the need for equitable participation: “You have to make sure that everyone’s involved, not just ‘leaders’ of various groups” (Community Tool Box undated)
- the potential disruption to the evaluation by participants’ personal circumstances (“illness, child care and relationship problems, getting the crops in, etc.”) (Community Tool Box undated)

One document states that “using a stakeholder model is only relevant if stakeholders are able to influence the programme or organization evaluated.” (Hansen 2005, citing Cameron 1986)

The following quote contains a fundamental criticism of empowerment evaluation:

Scriven (1997) has expressed concerns regarding its credibility as well as the relationship between internal and external evaluation. He believes “the credibility of a favorable evaluation done by an independent evaluator is obviously of much greater value to the staff with respect to external audiences than the issuance of favorable self-evaluation.” (Thiel & Feendy undated)

**Steps: Example 6**

The steps below are part of a framework “written to help the public health workforce gain a common understanding of evaluation concepts and to promote further integration of evaluation into the routine activities of public health organizations.” (Milstein et al. 2000) The framework is used for different purposes, e.g. guiding evaluations, clarifying programs, developing policies.

The steps outlined below are to be carried out in conjunction with the set of evaluation standards described in the section *Guiding Principles* (utility, feasibility, propriety, accuracy):

The standards provide sound guidelines to follow when having to decide among evaluation options. In particular, the standards help avoid creating an imbalanced evaluation (e.g., one that is accurate and feasible but not useful or one that would be useful and accurate but is unethical and therefore infeasible).
These standards can be applied both while planning an evaluation and throughout its implementation. (Milstein et al. 2000)

1. **Engage Stakeholders:** “After becoming involved, stakeholders help execute the other steps.” Key stakeholder groups are: “those involved in program operations, those affected by the program, and primary users of the evaluation.”

2. **Describe The Program:** e.g. mission, objectives, strategies, expected effects, resources, stage of development, “the logic model that displays how the entire program is supposed to work.”

3. **Focus the Evaluation Design:** e.g. consider evaluation “purpose, users, uses, questions, methods, and the agreements that summarize roles, responsibilities, budgets, and deliverables for those who will conduct the evaluation.”

4. **Gather Credible Evidence:** e.g. identify: stakeholders’ criteria for acceptable evidence, nature of indicators, sources, how to “gather and handle evidence.”

5. **Justify Conclusions:** Consider data “from a number of different stakeholder perspectives, to reach conclusions that are justified. Conclusions become justified when they are linked to the evidence gathered and are consistent with agreed-on values or standards set by stakeholders…the process of reaching justified conclusions involves four basic steps: (a) analysis/synthesis, to determine the findings; (b) interpretation, to determine what those findings mean; (c) judgments, to determine how the findings should be valued based on the selected standards; and (d) recommendations, to determine what claims, if any, are indicated.”

6. **Ensure Use and Share Lessons Learned:** “Some activities that promote use and dissemination include designing the evaluation from the start to achieve intended uses, preparing stakeholders for eventual use, providing continuous feedback to stakeholders, scheduling follow-up meetings with intended users to facilitate the transfer of conclusions into appropriate actions or decisions, and disseminating lessons learned to those who have a need or a right to know or an interest in the project.”

Figure 6 below is from Milstein et al. (2000).
Figure 6: Chronic Disease Centre’s Evaluation Framework (Milstein et al. 2000)

(Note: To view the Chronic Disease Centre’s evaluation framework, whose steps are listed above, visit the CDC Evaluation Working Group’s Web site at http://www.cdc.gov/eval/index.htm.)

Balanced Scorecard

The scorecard translates the vision and strategy of a business unit into objectives and measures in four different areas: the financial, customer, internal-business-process and learning and growth perspectives. The financial perspective identifies how the company wishes to be viewed by its shareholders. The customer perspective determines how the company wishes to be viewed by its customers. The internal-business-process perspective describes the business processes at which the company has to be particularly adept in order to satisfy its shareholders and customers. The organizational learning and growth perspective involves the changes and improvements which the company needs to realize if it is to make its vision come true. (Norreklit 2000)

The premise of the Balanced Scorecard (BSC) is that financial measurements are insufficient in the information age and that other perspectives are required to balance the financial perspective: “While competitive advantage was mainly driven by superior management of tangible assets in the industrial age, it is driven today by managing intangibles such as service, innovation, flexibility or...
knowledge.” (Speckbacher et al. 2003) According to the BSC, intangibles such as knowledge should not be translated into financial measures.

For each of the BSC’s four perspectives (financial, customer, business process, learning and growth), measurements are identified and data collected and analyzed. Although initially developed for a business context, the scorecard is now used in other settings such as health care and government. Its relationship to evaluation is that it “contains outcome measures and the performance drivers of outcomes, linked together in cause-and-effect relationships.” (Norreklit 2000)

Defining the BSC is not straightforward, given the “ambiguous definitions of the BSC concept in the theoretical literature.” (Speckbacher et al. 2003) Some documents identify cause and effect between the four perspectives as central to the BSC, for example: “The crux of the balanced scorecard is the linking together of the measures of the four areas in a causal chain which passes through all four perspectives...” (Norreklit 2000) Others however do not agree. For example, one document identifies three types of BSC, with the first type omitting cause and effect:

— Type I BSC: a specific multidimensional framework for strategic performance measurement that combines financial and non-financial strategic measures
— Type II BSC: a Type I BSC that additionally describes strategy by using cause-and-effect relationships
— Type III BSC: a Type II BSC that also implements strategy by defining objectives, action plans, results and connecting incentives with BSC (Speckbacher et al. 2003)

Figure 7 below is from the Balanced Scorecard Institute (b).
As with other evaluation approaches, documents cite many benefits from use of the BSC. For example, in one study, “The benefit of improving the alignment of strategic objectives with the actions actually undertaken was ticked most often.” (Speckbacher et al. 2003) Norreklit (2000) quotes some of the benefits of the BSC identified by Kaplan and Norton, the BSC’s originators:

— clarify and gain consensus about strategy;
— align departmental and personal goals to strategy;
— link strategic objectives to long-term targets and annual budgets;
— identify and align strategic initiatives; and
— obtain feedback to learn about and improve strategy...

However, Norreklit (2000) states that the BSC “is a hierarchical top-down model and that it is not easily rooted in a dynamic environment or in the organization.” The result is a gap between the planned strategy and the strategy actually carried out. Further, Norreklit refutes the basic assumption of Types II and III BSC, the causal chain, arguing that there is a logical but not a causal relationship among the four perspectives. For example, “Customer satisfaction does not necessarily yield good financial results.” As a result, “the balanced
scorecard makes invalid assumptions, which may lead to the anticipation of performance indicators which are faulty, resulting in sub-optimal performance.”

**Steps: Example 7**

A set of BSC implementation steps follow, from *Building & Implementing a Balanced Scorecard: Nine Steps to Success®* (Balanced Scorecard Institute a):

1. **Assessment:** assess organization’s mission and vision, challenges, enablers, and values; prepare a change management plan; conduct “a focused communications workshop to identify key messages, media outlets, timing, and messengers.”

2. **Strategy:** develop strategic results, strategic themes, and perspectives “to focus attention on customer needs and the organization’s value proposition.”

3. **Objectives:** decompose strategic elements from the first two steps “into Strategic Objectives, which are the basic building blocks of strategy and define the organization’s strategic intent. Objectives are first initiated and categorized on the Strategic Theme level, categorized by Perspective, linked in cause-effect linkages (Strategy Maps) for each Strategic Theme, and then later merged together to produce one set of Strategic Objectives for the entire organization.”

4. **Strategic Map:** formalize “the cause and effect linkages between the enterprise-wide Strategic Objectives” into a Strategy Map. “The previously constructed theme Strategy Maps are merged into an overall enterprise-wide Strategy Map that shows how the organization creates value for its customers and stakeholders.”

5. **Performance Measures:** develop performance measures “for each of the enterprise-wide Strategic Objectives. Leading and lagging measures are identified, expected targets and thresholds are established, and baseline and benchmarking data is developed.”

6. **Initiatives:** develop strategic initiatives to support objectives.

7. **Automation:** use “performance measurement software to get the right performance information to the right people at the right time…automation helps people make better decisions because it offers quick access to actual performance data.”

8. **Cascade:** “cascade” the enterprise-level scorecard “down into business and support unit scorecards, meaning the organizational level scorecard (the first Tier) is translated into business unit or support unit scorecards (the second Tier) and then later to team and individual scorecards (the third Tier). Cascading translates high-level strategy into lower-level objectives, measures, and operational details…Team and individual scorecards link day-to-day work with department goals and corporate vision…Performance measures are developed for all objectives at all organization levels… [and] become more operational and tactical…”
9. **Evaluation**: evaluate the completed scorecard: “Are our strategies working?” “Are we measuring the right things?” “Has our environment changed?” “Are we budgeting our money strategically?”

Figure 8 below, from the Balanced Scorecard Institute (a), summarizes the steps outlined above.

**Figure 8: Balanced Scorecard Steps (Balanced Scorecard Institute a)**

---

**Appreciative Inquiry**

While organizational development methods differ greatly depending on the purpose of the intervention and the organization’s population and context, many approaches tend to focus on identifying specific problems, analyzing possible causes and solutions to these problems, and devising a plan to resolve or eliminate the problems.

Appreciative Inquiry looks at organizational issues, challenges, and concerns in a significantly different way. Instead of focusing on problems, organizational members first
discover what is working particularly well in their organization. Then, instead of analyzing possible causes and solutions, they envision what it might be like if “the best of what is” occurred more frequently. Here participants engage in a dialogue concerning what is needed, in terms of both tasks and resources, to bring about the desired future. Finally, organization members implement their desired changes... (Coghlan et al. 2003)

Appreciative Inquiry (AI), designed as an organizational development approach and currently used in business and other organizational settings, has strong evaluation applications. Following are examples from Coghlan et al.’s list of “situations in which we believe Appreciative Inquiry has the most potential to contribute to evaluation practice”:

— Where there is a fear of or skepticism about evaluation
— With varied groups of stakeholders who know little about each other or the program being evaluated
— When relationships among individuals and groups have deteriorated and there is a sense of hopelessness
— When there is a desire to build evaluation capacity—to help others learn from evaluation practice
— To guide an evaluation’s design, development, and implementation as an overarching philosophy and framework

Regarding strengths of AI, Coghlan et al. (2003) suggest that AI can “increase participation in the evaluation, maximize the use of results, and build capacity for learning and change within organizations and communities.” Rogers and Fraser (2003) state that:

One of its major strengths is its fundamental recognition that an evaluation is an intervention that causes ripples in the life of an institution. Many evaluators operate without recognition of this fact, and they proceed with their assessments in ways that are insensitive to the impact their questions and investigations have on the life of an organization or institution.

They also state that, “Appreciative Inquiry can help to identify and make explicit areas of good performance and to communicate and institutionalize what is already known about good performance so that it gets continued or replicated.”

A common criticism of AI is that by focusing on the positive it ignores challenges, for example AI “risks encouraging avoidance of known problems, which is rarely an effective response.” (Rogers & Fraser 2003)

Coghlan et al. and Patton both dispute this criticism:

Appreciative Inquiry does address issues and problems, but from a different and often more constructive perspective: it reframes problem statements into a focus on strengths and successes. For example, rather than ask participants to list the problems their organization is facing, they are asked to explain what is going well, why it is going well, and what they want more of in the organization. In some Appreciative Inquiry efforts,
participants are also asked to state their specific wishes for the organization. This implicitly raises and addresses problems. (Coghlan et al. 2003)

The focus on appreciation can imply an unwillingness to look at weaknesses, problems, and things that are going wrong. Yet the cases in this volume provide strong evidence that problems and weaknesses can and do emerge in an appreciation-centered inquiry. Moreover, the cases even provide evidence that some problems and weaknesses can be easier to address and surface when evaluation takes an appreciative stance...

One can make the case that needs assessment is as unbalanced as Appreciative Inquiry, only in the opposite direction. Indeed, I might speculate that an Appreciative Inquiry evaluation is more likely to turn up problems and weaknesses (unmet needs) than a needs assessment is likely to turn up assets and strengths. Of course, in both instances, it all depends on how—and how well—it is done. (Patton 2003)

Rogers and Fraser (2003) raise a number of other criticisms of AI in addition to the potential ignoring of challenges, for example:

Appreciative Inquiry can be a useful and valuable technique in the right circumstances and when well implemented, but it is not always appropriate and it requires special skills and abilities to be done properly.

We are also concerned about the overemphasis on how perceptions affect reality without due concern for how reality can intrude on people’s lives. There are many aspects of lived reality that are not invented and exist despite our mental state; grinding poverty, gender inequality, violence, and disease are some examples.

One of our main concerns around Appreciative Inquiry is its potential to involve and include program staff (who are better able to participate in more of the processes) at the expense of program beneficiaries, and in this way to deepen and widen the gap between those who are already empowered and those who are not.

Steps: Example 8

The steps below describe the 4-D model of AI as outlined in Coghlan et al. (2003).

1. **Discovery (appreciate):** “consists of participants interviewing each other and sharing stories about their peak experiences,” e.g. “...what is it that you most value about yourself, your work, and your organization?” and “What three wishes do you have to enhance the health and vitality of your organization?” Groups develop an interview protocol based on the key themes arising out of the stories. “Using the new protocol, interviews are conducted with as many organization members as possible, ideally by the members themselves:

2. **Dream (envision results):** “envision themselves and their organization functioning at their best. Through various kinds of visualization and other creative exercises, participants think broadly and holistically about a desirable future.”

3. **Design (coconstruct the future):** “participants propose strategies, processes, and systems; make decisions; and develop collaborations that will create and support positive change. They develop provocative propositions or possibility and design statements that are concrete, detailed visions based on what was discovered about past successes.”

4. **Destiny (sustain the change):** implement visions of the Dream phase and propositions of the Design stage. “This phase is ongoing as participants continue to
implement changes, monitor their progress, and engage in new dialogue and Appreciative Inquiries.”

Figure 9 below summarizes a different AI approach, the 4-I model (from Coghlan et al. 2003, adapted from Watkins and Mohr 2001). According to Coghlan et al.:

The models, which are similar, have two major differences: (1) they use different language to describe the various phases, and (2) they present a different delineation of the phases. The 4-D model has a Destiny or Deliver phase that relates to implementation, while the 4-I model has an extra early planning step, Initiate, and embeds implementation into the Innovate phase.
Figure 9: Appreciative Inquiry 4-I Model (Coghlan et al. 2003, adapted from Watkins & Mohr 2001)

- **Initiate**
  - Introduce key stakeholders to AI theory and practice.
  - Create temporary project structures (sponsor team and core group), and educate sponsor team and core group in AI theory and practice.

- **Innovate**
  - Engage maximum possible number of organization members in conversations that enable exploration of and commitment to whatever actions, new roles, relationships, or design modifications (the social architecture of the organization) are seen as being important to support.
  - Implementation of the provocative propositions.
  - Implement the design changes using an AI-based progress review process.

- **Imagine**
  - Collate and share interview data and pull out themes (life-giving forces).
  - Develop provocative propositions (a grounded vision of the desired future).
  - Consensually validate provocative propositions with as many members of the system as possible.

- **Inquire**
  - Conduct generic interviews (this may also be done in the Initiate phase as part of the core group and sponsor team education).
  - Develop customized interview protocol; pilot and revise protocol (often this is the core group, with as much involvement by the steering committee as possible).

- **Interview**
  - Maximum possible number of client system members are interviewed.
Other Approaches

[There is an] increasingly rich array of developmental diversity within evaluation that includes participatory evaluation, collaborative evaluation, empowerment evaluation, inclusive evaluation, developmental evaluation, multivocal evaluation, learning-oriented evaluation, democratic evaluation, feminist evaluation, and now Appreciative Inquiry evaluation, or simply appreciative evaluation. (Patton 2003)

Some of the approaches listed in the quote above, and others not listed, are described in previous sections. A few more evaluation approaches are briefly described below.

**external**

This approach employs an external evaluator (a third party or person/organization not previously associated with the project being evaluated) to conduct the evaluation. (Public Safety and Emergency Preparedness Canada 2006)

A key strength of this approach is that an external evaluator may increase credibility of evaluation results. In addition, external evaluations may be less time consuming. External evaluation approaches can be combined with other approaches, for example goals based or utilization focused.

A key caution is that “research carried out by researchers who have not involved key stakeholders in the research process runs the risk of being irrelevant to the stakeholders, not understood by them, and/or ignored by them.” (Kahan & Goodstadt 2005) Along the same lines, external evaluations “can sometimes lead to missed information, or a misinterpretation of information.” (Farell et al. 2002)

**Kirkpatrick**

The approach developed by Kirkpatrick in the 1970s has four levels, but the literature describes the levels differently, in particular the last level. One description follows (from McEwen 2006, adapted from Gurksy 2000):

Four-level model to judge the quality, efficiency and effectiveness of programs:
- reaction – how participants feel
- learning – knowledge, skills and attitudes
- behavior – how behavior changes
- results – how processes have improved

According to Paul Bernthal (quoted in Abernathy 1999), “The simplicity and common sense of Kirkpatrick’s model imply that conducting an evaluation is a standardized, prepackaged process. But other options are not spelled out in the model.”

**CIPP**

CIPP stands for context, input, process, product and was developed by Stufflebeam in the 1960s. According to the Adult Literacy Accreditation Project:

The CIPP model of evaluation concentrates on:
— Context of the programme
— Input into the programme
— Process within the programme
— Product of the programme

The CIPP model proposed by Stufflebeam was also a reaction to Tyler's approach which concentrated on the final outcomes of a learning program. This model requires evaluators to study the effectiveness and feasibility of the initial policy decisions that gave rise to the programme as well as the programme's operation. In any evaluation according to this model the influence of institutional priorities, the impact of individual personalities and the importance of the prevailing political climate would be considered.

The value of the CIPP model is that it takes into consideration the whole context surrounding a programme. Its disadvantage may be the time necessary to carry it out.

Summary Note on Evaluation Approaches
Choosing the best evaluation approach is not always easy. Rogers and Fraser (2003) suggest criteria for selecting an approach: plausibility (does it appear likely that the evaluation approach will be helpful); practicality (is it possible to implement the approach); and evidence that the approach works.

In practice, evaluations rarely follow only one evaluation approach. For example, results-focused approaches do not preclude working collaboratively with a number of stakeholder groups and collaborative approaches can easily adopt a utilization-focused approach. This section ends with the following comment on assessing and selecting individual evaluation approaches:

...we do not expect any single evaluation approach or model to be sufficient for a given evaluation and do not evaluate them in terms of their comprehensiveness. We see different evaluation approaches as components in our repertoire that can be combined as needed, not as mutually exclusive commercial products that should be adopted or rejected in their entirety. Nor do we expect any approach to be suitable for all programs, all evaluation purposes, all countries, and all organizations. (Rogers & Fraser 2003)
EVALUATION IN PRACTICE

Considering potential benefits, challenges and suggestions for success will increase the likelihood of an evaluation framework successfully guiding an evaluation. A summary of each of these topics, based on a review of the literature, follows.

Benefits

While at first one may just think about the use of findings and recommendations, it is clear that there are other aspects of an evaluation that can also be used. The evaluation process itself can be used as a forum for dialogue and learning. Ideas and generalizations produced by the evaluations can be used. Sometimes even the fact that an evaluation is undertaken is used to legitimize the seriousness of the management style or the policy instrument. The focus and the measures used in the evaluation can be used to change implementation. If, for example, transparency is one focus of an evaluation, some part of the implementation process which nobody has prioritized earlier may become more open by making documents available on the Internet, or the opposite could happen. Finally the design of the evaluation can be used by other evaluators or by groups emphasizing different views... (Mickwitz 2003)

A list of some of the most commonly mentioned benefits resulting from conducting evaluations follows.

- supports accountability
- builds capacity
- supports communication, e.g. by providing a historical record or current description and status of the policy or program
- increases understanding, e.g. re.: intended and unintended results of the policy or program; why a program or policy is or isn’t successful; how to address challenges; evaluation
- provides information for decision making
- increases improvement, e.g. of processes, activities, outcomes

Challenges

Without explicit evaluation guidelines, insensitive and inappropriate—and perhaps exploitative and harmful—evaluative research practices may result. It is not uncommon in evaluation research to implement outcome measures that have only been validated and normed on White, middle-class samples. Moreover, many national evaluations have focused solely on outcome rather than process and have paid little attention to within-group variation, opting instead to examine between-group differences (McLoyd and Randolph 1985). Such research practices not only ignore the variability and diversity within cultural groups but often rely on deficit-based perspectives where differences are conceptualized as “the genetic or the cultural inferiority of ethnic/racial groups relative to the White mainstream standard” (Garcia Coll et al. 1996, 1893). Much of the extant evaluation research assessing minority family outcomes continues to be guided by theoretical or conceptual models that do not place the interaction of social class, culture, ethnicity, and race at the core, thus minimizing the effects of social inequality and

A brief outline of key categories of challenges identified in the literature follows.

- **context**: e.g.: pressure from external forces to conduct only certain types of evaluations or use one method over another; a culture that does not value evaluation or views performance management as adequate to fulfill evaluation purposes
- **resources**: e.g. lack of time, funding, energy, personnel, skills
- **measurement**: e.g.: inappropriate indicators; emphasis on one kind of measurement over another such as exclusive use of quantitative measures or exclusive use of qualitative measures; the easily measured drives policy or program decisions rather than the program purpose driving what gets measured
- **data collection and analysis**: e.g.: inappropriate, inaccurate or insufficient data; misinterpretation of data; failing to take into account non-program variables when assessing program outcomes
- **sensitivity**: e.g. cultural and other kinds of insensitivity when dealing with evaluation participants
- **imbalance**: e.g.: only outcome evaluations are conducted or only process evaluations; long-term outcomes are ignored; focus is on deficits with assets ignored; financial costs are emphasized over human costs; not all stakeholder groups have the same opportunity to participate (“Different groups have different resources, such as members, money, personnel or intellectual capabilities, which affect their potential to participate.” [Mickwitz 2003])
- **participation**: e.g. engaging stakeholders; only leaders from various stakeholder groups participate
- **follow up**: evaluation results are ignored
- **planning**: insufficient thought is given to issues such as design, methods, stakeholder participation
- **power**: e.g.: managers have more power than front line workers; staff have more power than program participants; some program participants may have more power than others due to issues such as class, gender, age, race
- **complexity**: e.g. when multiple sites are involved

**Suggestions for Success**

The literature provides many suggestions for increasing the probability of a successful evaluation. One set of suggestions follows, from the *Good Practice Guide* (Department of Health 2002).

- View the evaluation as an integral part of the project and plan for it at the outset. The evaluation should be costed and resourced as part of the project.
- Secure commitment from senior managers within the organisation...
- Involve all key stakeholders in its planning and execution...
- Develop relevant criteria and indicators to assess project outcomes from the outset of the project.
- Put in place mechanisms to enable monitoring and measurement of progress.
- Foster a learning environment to ensure lessons are heeded...
— The potential value of an evaluation will only be realised when action is taken on the findings and recommendations emanating from it. Processes are needed to ensure that this happens...
— As with all aspects of the business case, the Evaluation Plan should be a live document. It should be kept under constant review. The existence of an Evaluation Plan allows changes in the project objectives and other important parameters to be explicitly noted...
— Roles and responsibilities in analysing the data and writing the report should be clearly defined in the Evaluation Plan...
— The findings of the evaluation should be validated. It is good practice to share drafts of the report with those who have provided the information (i.e., the study participants). It is important to ensure their views have been fairly and accurately taken into account...

While the above list emphasizes a stakeholder approach, the summary below, from the Treasury Board of Canada’s Evaluation Policy (2001) comes from a more “professional” stance:

The success of evaluation depends on a number of important factors. It requires clarity of roles, application of sound standards, ongoing support for rigorous, professional practice, and developing a conducive environment where managers embed the discipline of evaluation into their work. The organisational positioning of evaluation should reflect the unique needs of the department. Evaluation discipline should be used in synergy with other management tools to improve decision-making. Heads of evaluation should work to ensure that evaluation in their organisation is healthy in respect of all these factors.

Documents provide many other suggestions than those mentioned above. One common suggestion is to review processes as well as outcomes. Scriven (2007) writes:

Here are four other examples of why process is an essential element in program evaluation, despite the common tendency in much evaluation to place almost the entire emphasis on outcomes:
(i) gender or racial prejudice in selection/promotion/treatment of staff is an unethical practice that must be checked for and comes under process
(ii) in evaluating health programs that involve medication, “adherence” or “implementation fidelity” means following the prescribed regimen including drug dosage, and it is often vitally important to determine the degree to which this is occurring—which is also a process consideration...;
(iii) the process may be where the value lies—writing poetry in the creative writing class may be a good thing to do in itself, not because of some later outcomes (same for having fun, in kindergarten at least; painting; and marching to protest war, even if it doesn’t succeed);
(iv) the treatment of human subjects must meet federal, state, and other ethical standards, and an evaluator cannot avoid the responsibility for checking that they are met.

Another common suggestion is to use a variety of methods and designs and to be flexible in their application:
The importance of a range of research methods is paramount. Policy evaluation uses quantitative and qualitative methods, experimental and non-experimental designs, descriptive and experiential methods, theory based approaches, research synthesis methods, and economic evaluation methods. It privileges no single method of inquiry and acknowledges the complementary potential of different research methods. The methods used in policy evaluation and analysis are usually driven by the substantive issues at hand rather than a priori preferences (Greene, Benjamin and Goodyear, 2001). (Government Social Research Unit 2007)

Models with a primary focus on accountability may not always provide the best fit for the various purposes of educational evaluations. Therefore, it is important to employ a balance of evaluation strategies to ensure that the evaluation methods align with the project goals. Reliance on any single method provides incomplete evidence of the effectiveness of a program, a method, an innovation, or a product. A developmental model for evaluation is suitable for many educational settings and purposes. In this approach, evaluations are tailored to the specific needs of the program or project being evaluated. This allows for contextualization of the evaluation which adds to the validity of inferences that can be drawn from the findings and any recommendations for change. (Thiel & Feendy undated)

We’re in the information age, and the thinking is that our evaluation methods should reflect that. Certainly, the best strategy considers what’s best for an organization, its employees, and its customers. It also helps us look forward to future value and backward to past results. And that suggests a balanced view that considers hard- and soft-skill performance gauges, tangible and intangible benefits, and short- and long-term results in training valuation (Abernathy 1999).

Davies (1999) stresses, among other things, the importance of not forcing “evaluative enterprise in general” on people because of the negative impacts that may result such as unreliable information. Davies also talks about agreeing in advance on topics such as measures and use of information.

One aspect to take into account of when planning evaluations relates to a point made by Patton (2003):

From a systems perspective, because programs and projects are usually embedded in larger organizational contexts, improving programs and projects may be linked to and even dependent on changing the organizations of which they are a part.

Another consideration is the importance of relationship building and working collaboratively to ensure cultural relevance and evaluation success, as discussed in Evaluating from the Outside: Conducting Cross-Cultural Evaluation Research on an American Indian Reservation (Letiecq and Bailey 2004):

Perhaps one of the more salient lessons learned has been the importance of relationship building and the need to reaffirm such relationships often. The historical injustices experienced by tribal communities and the misuse of tribal knowledge requires constant dialogue and frequent meetings to ensure cross-cultural understanding and appropriateness (e.g., “What is the culturally appropriate way to greet an elder?” “Is the wording of this item acceptable on the interview schedule?”). Such dialogue can be time
consuming and is not often valued by academic culture. Yet to be effective as outside evaluators, this time spent on relationship building is critically important and may determine the success or failure of a project...

It is critical—particularly in cross-cultural research and evaluation—to jointly determine a research protocol or set of standards to be followed throughout the course of the evaluation. Determining who the stakeholders are, how data will be used, and to what end should be a collaborative process where the tribal community has the final say (Fisher and Ball 2002). Again, from an academic perspective, this lack of control over data can make analysis and dissemination of data challenging, and does not always fit well with academe’s “publish or perish” cultural norm. Nonetheless, partnering with tribal nations to conduct research and evaluation can be an advantageous endeavor, as the cross-fertilization of two or more cultures—blending ways of knowing and assessments of success—can produce synergistic, creative, and mutually rewarding projects that advance our understanding of the strengths and needs of diverse communities.

**Incorporating Evaluation into an Organization**

Several of the documents reviewed provide information regarding incorporating evaluation into an organization. For example, Stufflebeam’s *Institutionalizing Evaluation Checklist* (2002) lists 18 points. Four are provided below as examples:

- Since evaluation systems are **context dependent**, take into account constituents’ needs, wants, and expectations plus other variables such as pertinent societal values, customs, and mores; relevant laws and statutes; economic dynamics; political forces; media interests; pertinent substantive criteria; organizational mission, goals, and priorities; organizational governance, management, protocols, and operating routines; and the organization’s history and current challenges.
- Define and pursue clear, appropriate **evaluation purposes**: improvement, accountability, organizational learning, and dissemination.
- Engage and support a capable, credible **evaluation team** possessing expertise in field work, group process, interviewing, measurement, statistics, surveys, cost analysis, values analysis, policy analysis, public speaking, writing, editing, computers, communications technology, and project management plus diverse substantive, gender, and ethnic backgrounds, as needed.
- Supply the evaluation effort with sufficient **funds, facilities, equipment, services, software, and technical support**.

The Saskatchewan Department of Community Resources and Employment’s 2003 document provides guidelines for implementing a systematic approach to evaluation (and other research) into Social Services. The document emphasizes the importance of a culture which values continuous learning and improvement. Examples of this document’s recommendations for systematising evaluation in the Ministry include:

- review new programs after a year of operation
- conduct effectiveness evaluations on new programs by the end of their first three years
- require all new programs to have plans for monitoring and evaluation
- make evaluation decisions in synch with planning and budget cycles.

The Treasury Board of Canada’s *Evaluation Policy* (2001) also recommends that “departments should embed the discipline of evaluation into the lifecycle management of policies, programs and initiatives.” Further,

> It is government policy that departments and agencies embed evaluation into their management practices:
> — to help design policies, programs, and initiatives that clearly define expected results and that embody sound performance measurement, reporting and accountability provisions at their outset; and
> — to help assess in a rigorous and objective manner the results of government policies, programs, and initiatives, including their impacts both intended and unintended, and alternative ways of achieving expected results.

The document *Evaluation Policy and Framework* (Saskatchewan Post-Secondary Education and Skills Training 1998) contains a useful table of roles and responsibilities which could be adapted to the Ministry of Education. This table is presented below. One key informant suggested as part of the adaptation defining the category of “senior management” further, for example which tasks were the specific responsibilities of executive directors, the leadership forum, deputy minister’s office, or executive committee.

*Table 3: Roles and Responsibilities (Saskatchewan Post-Secondary Education and Skills Training 1998)*

<table>
<thead>
<tr>
<th>ROLES and RESPONSIBILITIES</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Senior Management</strong></td>
</tr>
<tr>
<td>• Approve evaluation policy and framework.</td>
</tr>
<tr>
<td>• Ensure availability of adequate resources (human and financial) to support evaluation.</td>
</tr>
<tr>
<td>• Approve the department's annual evaluation schedule.</td>
</tr>
<tr>
<td>• Approve detailed evaluation plans.</td>
</tr>
<tr>
<td>• Review evaluation reports.</td>
</tr>
<tr>
<td>• Use evaluation results as a basis for decision-making and policy formulation.</td>
</tr>
<tr>
<td><strong>The Policy and Evaluation Branch, PSEST</strong></td>
</tr>
<tr>
<td>• Facilitate development and maintenance of evaluation policies and strategies.</td>
</tr>
<tr>
<td>• Facilitate development of and coordinate the department’s annual evaluation schedule.</td>
</tr>
<tr>
<td>• Ensure evaluation expertise is available.</td>
</tr>
<tr>
<td>• Provide advice, guidance, co-ordination, and project management as needed.</td>
</tr>
<tr>
<td>• Ensure the maintenance of recognized evaluation standards.</td>
</tr>
<tr>
<td>• Participate on evaluation working teams.</td>
</tr>
<tr>
<td><strong>Policy, Program and Service Areas</strong></td>
</tr>
<tr>
<td>• Develop and execute evaluation activities (planning, annual self-assessment, and detailed evaluations) in a timely manner.</td>
</tr>
<tr>
<td>• Ensure inclusion of evaluation plans in all new initiatives.</td>
</tr>
<tr>
<td>• Participate on evaluation working teams.</td>
</tr>
<tr>
<td>• Use evaluation results as a basis for decision-making.</td>
</tr>
<tr>
<td>• Ensure follow-up on the recommendations.</td>
</tr>
<tr>
<td><strong>Evaluation Working Team (EWT)</strong></td>
</tr>
<tr>
<td>• Ensure development and execution of the detailed evaluation for a specific policy, program or service in a timely manner.</td>
</tr>
<tr>
<td>• Participate in development of Requests for Proposals (RFPs) for evaluation</td>
</tr>
<tr>
<td>Contractors</td>
</tr>
<tr>
<td>----------------------------------------------</td>
</tr>
<tr>
<td>• Select/recommend contractors to implement the evaluation.</td>
</tr>
<tr>
<td>• Oversee evaluation contractors.</td>
</tr>
<tr>
<td>• Analyze the evaluation results and conclusions.</td>
</tr>
<tr>
<td>• Develop recommendations.</td>
</tr>
<tr>
<td>• Finalize report.</td>
</tr>
<tr>
<td>• Participate in development of follow-up plans.</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
REFERENCES


Adult Literacy Accreditation Project. WIT (Waterford Institute of Technology) and NALA (National Adult Literacy Agency) Dublin, Ireland. (Undated.)


Balanced Scorecard Institute (a). *Building & Implementing a Balanced Scorecard: Nine Steps to Success®*. (Undated.)


Balanced Scorecard Institute (b). *What is the Balanced Scorecard?*. (Undated.)


Canadian Evaluation Society. *CES Guidelines For Ethical Conduct*. (Undated.)

Bibliographic reference: [website: http://evaluationcanada.ca/site.cgi?s=5&ss=4&_lang=en](http://evaluationcanada.ca/site.cgi?s=5&ss=4&_lang=en)


Community Tool Box (University of Kansas). *Chapter 36: Introduction to Evaluation*. (Undated.)

Bibliographic reference: [website: http://ctb.ku.edu/](http://ctb.ku.edu/)


Bibliographic reference: [website: www.wmich.edu/evalctr/checklists](http://www.wmich.edu/evalctr/checklists)

Hunter, Heather. Email communication, 2008.


Saskatchewan Industry and Resources (Planning and Evaluation Branch (P&E) Revenue and Funding Services). *Business Process Improvement (BPI) And Program Review Framework.* 2004


website: www.wmich.edu/evalctr/checklists


website: www.wmich.edu/evalctr/checklists


website: http://education.alberta.ca/apps/aisi/literature/pdfs/Project_Evaluation_UniversityofLethbridge.pdf


website: http://education.alberta.ca/media/325660/suppdoc9.pdf


website: http://www.thcu.ca/infoandresources/publications/EVALMaster.Workbook.v3.6.08.15.07.pdf


website: http://www.wkkf.org/Pubs/Tools/Evaluation/Pub3669.pdf

